Towards quantized number theory: spectral operators and an asymmetric criterion for the Riemann hypothesis

Michel L. Lapidus

Department of Mathematics, University of California, 900 University Avenue, Riverside, CA 92521-0135, USA

This research expository article not only contains a survey of earlier work but also contains a main new result, which we first describe. Given $c \geq 0$, the spectral operator $a = a_c$ can be thought of intuitively as the operator which sends the geometry onto the spectrum of a fractal string of dimension not exceeding $c$. Rigorously, it turns out to coincide with a suitable quantization of the Riemann zeta function $\zeta = \zeta(s)$: $a = \zeta(\partial)$, where $\partial = \partial_c$ is the infinitesimal shift of the real line acting on the weighted Hilbert space $L^2(\mathbb{R}, e^{-2ct} dt)$. In this paper, we establish a new asymmetric criterion for the Riemann hypothesis (RH), expressed in terms of the invertibility of the spectral operator for all values of the dimension parameter $c \in (0, \frac{1}{2})$ (i.e. for all $c$ in the left half of the critical interval $(0, 1)$). This corresponds (conditionally) to a mathematical (and perhaps also, physical) ‘phase transition’ occurring in the midfractal case when $c = \frac{1}{2}$. Both the universality and the non-universality of $\zeta = \zeta(s)$ in the right (resp., left) critical strip ($\frac{1}{2} < \text{Re}(s) < 1$) (resp., $0 < \text{Re}(s) < \frac{1}{2}$)) play a key role in this context. These new results are presented here. We also briefly discuss earlier joint work on the complex dimensions of fractal strings, and we survey earlier related work of the author with Maier and with Herichi, respectively, in which were established symmetric criteria for the RH, expressed, respectively, in terms of a family of natural inverse spectral problems for fractal strings of Minkowski dimension $D \in (0, 1)$, with $D \neq \frac{1}{2}$, and of the quasi-invertibility of the family of spectral operators $a_c$ (with $c \in (0, 1)$, $c \neq \frac{1}{2}$).
1. Introduction

Our main goal in this paper is to first briefly explain (in §2) the general ideas and results concerning fractal strings and their complex dimensions which led to the intimate connections (discussed in §3) between the vibrations of fractal strings and the Riemann zeta function \( \zeta = \zeta(s) \) and, then, the Riemann hypothesis (RH), discovered in the early 1990s by the author and his co-workers [1–4]. The resulting geometric interpretation of the critical strip \( 0 \leq \Re(s) \leq 1 \) for \( \zeta = \zeta(s) \) was made entirely rigorous by means of the mathematical theory of complex dimensions developed by the author and Machiel van Frankenhuysen in a series of research monographs [5–7] (and papers). Owing to space limitations, we will not be able to do justice to this theory, for which the interested reader can refer to the latest book in the above series [7], and many relevant references therein.

In §6.3.1 and §6.3.2 of [6,7] was introduced, at the semi-heuristic level, the notion of spectral operator \( a = a_c \), which sends the geometry onto the spectrum of a fractal string. Intuitively, the parameter \( c \) plays here the role of the upper bound for the Minkowski dimensions of the fractal strings on which \( a = a_c \) acts. In a forthcoming book [8], and in an associated series of papers [9–12], Hafedh Herichi and the author have developed a mathematical theory of the spectral operator, within a rigorous functional analytic framework. In particular, they have precisely determined its spectrum as well as that of the closely related infinitesimal shift (of the real line) \( \partial = \partial_c \), viewed as the first-order differential operator \( d/dt \) acting on a suitable Hilbert space of functions, the weighted \( L^2 \)-space \( \mathbb{H}_c = L^2(\mathbb{R}, e^{-2ct} dt) \), where \( c \in \mathbb{R} \) is arbitrary. It turns out that \( a_c = \zeta(\partial_c) \) or, more briefly, \( a = \zeta(\partial) \), in the sense of the functional calculus for the unbounded normal operator \( \partial = \partial_c \). In this manner, the spectral operator \( a = a_c \) can be viewed as a suitable quantization of the Riemann zeta function. Actually, in [8] (see also [12]), a quantized (i.e. an operator-valued) Dirichlet series, Euler product representation and ‘analytic continuation’ of \( a_c = \zeta(\partial_c) \) are obtained (for suitable values of \( c \) and/or under appropriate assumptions).

In this context, the question of the invertibility of the spectral operator is a key question. It is shown in [8] (see also [9,10]) that \( a = a_c \) is ‘quasi-invertible’ for all \( c \in (0, 1) \), \( c \neq \frac{1}{2} \) (equivalently, for all \( c \in (0, \frac{1}{2}) \) or for all \( c \in (\frac{1}{2}, 1) \), respectively) if and only if the RH is true (i.e. iff \( \zeta(s) = 0 \) with \( 0 < \Re(s) < 1 \) implies that \( \Re(s) = \frac{1}{2} \)). This result provides a natural operator theoretic counterpart of the result of the author and Helmut Maier in [3,4] showing that a suitable inverse spectral problem, (ISP)\(_D\), for fractal strings (of dimension \( D \)) is true in all possible dimensions \( D \) in \((0, 1)\), \( D \neq \frac{1}{2} \) (or, equivalently, due to the functional equation of \( \zeta \), for all \( D \in (0, \frac{1}{2}) \) or else for all \( D \in (\frac{1}{2}, 1) \), respectively) if and only if the RH is true.

In [8] (see also [9,10]), the spectrum of \( a = a_c \) is shown to obey several ‘mathematical phase transitions’; one conditionally (i.e. if RH holds), in the midfractal case (in the sense of [13–15]) when \( D = \frac{1}{2} \), and another one unconditionally (i.e. independently of the truth of RH), at \( D = 1 \). The first phase transition is intimately connected with the universality of the Riemann zeta function [8,11]. A brief exposition of some of the most relevant results of [8–12] is provided in §4.

The spectral reformulation of the RH obtained in [4], revisited and extended in [5–7] in the light of the mathematical theory of complex dimensions, and given a rigorous operator theoretic version in [8–10], is a symmetric criterion for RH. Indeed, owing to the functional equation satisfied by \( \zeta = \zeta(s) \) (and thus connecting \( \zeta(s) \) and \( \zeta(1-s) \)), it can be equivalently formulated for all values of the underlying parameter \( (D \text{ in } [3,4] \text{ and } [5–7] \text{ or } c \text{ in } [8–12]) \). By contrast, the new criterion for the RH obtained in this paper is asymmetric, in the sense that it is only stated (and valid) for all values of the underlying parameter \( c \) in the open interval \((0, \frac{1}{2})\). In fact, its counterpart for all values of \( c \) in the symmetric interval \((\frac{1}{2}, 1)\) cannot be true, due to the universality of the Riemann zeta function \( \zeta = \zeta(s) \) in the right critical strip \( \frac{1}{2} < \Re(s) < 1 \). More specifically, the main new result of this paper is the following (see §5 and theorem 1.1): the spectral operator \( a = a_c \) is invertible (in the usual sense of the invertibility of an unbounded operator) for all values of \( c \) in \((0, \frac{1}{2})\) if and only if the RH is true. Again, this new asymmetric spectral reformulation of RH does not have any analogue for the interval \((\frac{1}{2}, 1)\) because the spectral operator is not invertible for any value of \( c \) in \((\frac{1}{2}, 1)\). This raises the question of finding an appropriate physical
interpretation for this asymmetry, which amounts to finding ‘the’ physical origin of the phase transition (conditionally) occurring in the midfractal case where \( c = \frac{1}{2} \).

In the previous context of the results of [3,4], early speculations were made by Lapidus [14,15] to find a physical interpretation of the corresponding mathematical phase transition at \( D = \frac{1}{2} \) in terms of Wilson’s notion of complex dimension within the realm of his theory of phase transitions and critical phenomena in quantum statistical physics and quantum field theory [16]. We leave it to the interested reader to find his or her physical interpretation of the asymmetric criterion for RH obtained in this paper. Our expectation is that, physically, this asymmetry or phase transition finds its source in the symmetry breaking associated with an appropriate (and possibly yet to be discovered) quantum field theory. (Unlike in [17,18] or [19, §V.11], it would occur at \( \frac{1}{2} \) rather than at 1, the pole of \( \zeta \).) Presumably, this quantum field theory should be intimately associated with number theory (or ‘arithmetic’), geometry, dynamics and spectral theory, and probably formulated in the spirit of the conjectural picture for the generalized Riemann hypothesis (GRH) proposed in [20], in terms of a Riemann flow on the moduli space of quantized fractal strings (called ‘fractal membranes’) and the associated flows of zeta functions (or partition functions) and of their zeros. However, we invite readers to follow their own intuition and to be led into unknown mathematical and physical territory, wherever their own imagination will guide them.

Finally, we stress that the main goal of this paper (besides providing in §§2–4 a brief survey of earlier relevant work which partly motivates it) is to obtain the following theorem, in §5. (See theorem 5.4 and the comments surrounding it for additional information.) Let \( b = aa^* = a^*a \) be the non-negative self-adjoint operator naturally associated with the normal operator \( a \). For notational simplicity, we write \( a \) and \( b \) instead of \( a_\mathcal{L} \) and \( b_\mathcal{L} \), respectively. Note that, for \( c \leq 1 \), both \( a \) and \( b \) are unbounded linear operators (according to the results of [8–11] discussed in §4).

**Theorem 1.1 (Asymmetric criterion for RH).** Each of the following statements is equivalent to the RH):

(i) For every \( c \in (0, \frac{1}{2}) \), the spectral operator \( a \) is invertible.

(ii) For every \( c \in (0, \frac{1}{2}) \), the unbounded operator \( b \) is invertible.

(iii) For every \( c \in (0, \frac{1}{2}) \), the non-negative self-adjoint operator \( b \) is bounded away from zero.

For some main notation and terminology, see the glossary.

2. **Geometry and spectra of fractal strings**

In this section, we give a brief overview of the theory of fractal strings and their complex dimensions [7].

(a) **Geometric zeta function and Minkowski dimension**

A *fractal string* is a one-dimensional drum with fractal boundary. Hence, it can be realized geometrically as a bounded open subset of the real line: \( \Omega \subset \mathbb{R} \). As is well known, \( \Omega \) is then the disjoint union of its connected components (open intervals). Therefore, we can write it as a countable union of disjoint (bounded) open intervals \( I_j, j = 1, 2, \ldots \): \( \Omega = \bigcup_{j=1}^{\infty} I_j \). Denoting by \( \ell_j \) the length of \( I_j \) (\( \ell_j = |I_j| \)), we may assume without loss of generality that \( \mathcal{L} = \{\ell_j\}_{j=1}^{\infty} \) is non-increasing and (if \( \mathcal{L} \) is infinite, which will always be assumed from now on, in order to avoid trivial special cases) that \( \ell_j \downarrow 0 \) as \( j \to \infty \). Since all of the key notions considered in fractal string theory are independent of the geometric realization \( \Omega \), we may consider that a fractal string is equivalently given by the sequence of lengths (or *scales*) \( \mathcal{L} = \{\ell_j\}_{j=1}^{\infty} \), with the distinct lengths written in non-increasing order and according to their multiplicities.

The geometric point of view, however, is essential to motivate and define the spectral problem associated with \( \mathcal{L} \), as well as to connect and contrast the theory of fractal strings with the higher-dimensional geometric and spectral theory of fractal drums.
To a given fractal string $\mathcal{L} = \{\ell_j\}_{j=1}^{\infty}$ (or $\Omega \subset \mathbb{R}$), we associate the geometric zeta function of $\mathcal{L}$, denoted by $\xi_{\mathcal{L}}$, and defined (for $s \in \mathbb{C}$ with $\text{Re}(s)$ large enough) by $\xi_{\mathcal{L}}(s) = \sum_{j=1}^{\infty} \ell_j^s$. As it turns out, $\xi_{\mathcal{L}}$ is holomorphic in the open, right half-plane $I := \{\text{Re}(s) > D\}$ (the set of all $s \in \mathbb{C}$ such that $\text{Re}(s) > D$), and $I$ is the largest such half-plane. Here, the unique extended real number $D \in \mathbb{R} \cup \{\pm \infty\}$ is called the abscissa of convergence of $\xi_{\mathcal{L}}$: $D = \inf\{\alpha \in \mathbb{R} : \sum_{j=1}^{\infty} \ell_j^\alpha < \infty\}$; e.g. [21,22].

Earlier ([14,15], [7,§1.2]), the author observed that $D$ coincides with the Minkowski dimension of $\mathcal{L}$ (i.e. of the boundary $\partial \Omega$ of any of its geometric realizations $\Omega$): $D = D_M$. Recall that this notion of fractal dimension (often also called ‘box dimension’ in the applied literature), which is very different mathematically and physically from the classic Hausdorff dimension [13], can be defined as follows: given $\varepsilon > 0$, let $\Omega_{\varepsilon} = \{x \in \Omega : \text{dists}(x, \partial \Omega) < \varepsilon\}$ denote the $\varepsilon$-neighbourhood (or tubular neighbourhood) of $\Omega$, and let $V(\varepsilon) = \|\Omega_{\varepsilon}\|$. Then, given $d \geq 0$, let $M^*_d$ (resp., $M_{s,d}$) denote the upper (resp., lower) $d$th dimensional upper (resp., lower) Minkowski content of $\mathcal{L}$. Clearly, we always have $0 \leq M_{s,d} \leq M^*_d \leq \infty$.

The Minkowski dimension of $\mathcal{L}$ (or, equivalently, of its boundary $\partial \Omega$), denoted by $D_M = D_M(\mathcal{L})$, is then given by $D_M := \inf\{d \geq 0 : M^*_d < \infty\} = \sup\{d \geq 0 : M_{s,d} = +\infty\}$. In fact, $D_M$ is the unique real number, such that $M^*_d = +\infty$ for $d < D_M$ and $M_{s,d} = 0$ for $d > D_M$ (e.g. [7,23–26]). Physically, it can be assumed as a critical exponent as it gives information about the behaviour of $V(\varepsilon)$ as a power of $\varepsilon$. In the sequel, for brevity and in light of the aforementioned identity $D = D_M$, we will use $D$ instead of $D_M$ to denote the Minkowski dimension of $\mathcal{L}$. A fractal string $\mathcal{L}$ (or, equivalently, its boundary $\partial \Omega$) is said to be Minkowski non-degenerate if $0 < M_s(\Omega), M^* < \infty$, where $M^*_d := M^*_D$ and $M_{s,d} := M_{s,D}$ denote, respectively, the upper and the lower Minkowski content of $\mathcal{L}$ (i.e. of $\partial \Omega$). In addition, if $M_s = M^*_d$, we denote by $M$ this common value and state that $\mathcal{L}$ (i.e. $\partial \Omega$) is Minkowski measurable and has Minkowski content $M$. Hence, $\mathcal{L}$ is Minkowski measurable with Minkowski content $M$ if and only if the following limit exists in $(0, +\infty)$ and $\lim_{\varepsilon \to 0^+} V(\varepsilon)/\varepsilon^{1-D} = M$.

(b) Geometric interpretation of the critical strip

For a fractal string, we always have $0 \leq D = D_M \leq 1$; that is, $D = D_M$ lies in the (closed) critical interval $[0,1]$, where (in the terminology of [13], later adopted in [1–12,14,15,20,27–31]) the case $D = 0$ (resp., $D = 1$) is referred to as the least (resp., most) fractal case, while the case when $D = \frac{1}{2}$ is called the midfractal case. The latter midfractal case $D = \frac{1}{2}$ plays a key role in the theory, in connection with the Riemann zeta function and the RH. Recall that, according to the celebrated RH [32] (translated in [33, appendix]), the critical zeros (also called the non-trivial zeros) of the Riemann zeta function $\zeta = \zeta(s)$, that is, the zeros of $\zeta$ located in the critical strip $0 < \text{Re}(s) < 1$, all lie on the critical line $\{\text{Re}(s) = \frac{1}{2}\}$. Equivalently, $\zeta(s) = 0$ with $0 \leq \text{Re}(s) \leq 1$ implies that $\text{Re}(s) = \frac{1}{2}$.

Indeed, towards the very beginning of the mathematical theory of fractal strings [1–4,14,15,30], one conjectured the existence of a notion of complex dimension such that the left-hand side of the (closed) critical strip $0 \leq \text{Re}(s) \leq 1$ (the vertical line $\{\text{Re}(s) = 0\}$) corresponds to the least fractal (or non-fractal) case $D = 0$, the right-hand side (the vertical line $\{\text{Re}(s) = 1\}$) corresponds to the most fractal case $D = 1$, while the middle of the critical strip (the critical line $\{\text{Re}(s) = \frac{1}{2}\}$) corresponds to the midfractal case $D = \frac{1}{2}$. The answer turned out to be positive, as was intuitively clear after the works in [1,2,14,15] and, especially, in [3,4], and was fully and rigorously justified in the mathematical theory of complex dimensions developed in [5–7] (as well as in earlier and later papers by the same authors). We refer to the research monograph [7] for a complete exposition of the theory of complex dimensions of fractal strings. We will briefly introduce it in §2d, but, first, we explain the origins of the connections between the Riemann zeta function and fractal strings, viewed as vibrating objects.
(c) From the geometry to the spectrum of a fractal string, via $\zeta = \zeta(s)$

A fractal string, $\Omega = \bigcup_{j=1}^\infty I_j$, can be viewed as countably many ordinary strings (the intervals $I_j$ of lengths $\ell_j$), vibrating independently of one another. Equivalently, the strings $I_j$ can be thought of as the strings of a fractal harp. The (normalized frequency) spectrum $\sigma(L)$ of the fractal string $L$ is given by the union (counting multiplicities) of the spectra of the intervals $I_j : \sigma(L) = \bigcup_{j=1}^\infty \sigma(I_j)$.

Note that the fundamental frequency of the $j$th string $I_j$ is equal to $\ell_j^{-1}$; so that, for each fixed integer $j \geq 1$, we have $\sigma(I_j) = \{n \cdot \ell_j^{-1} : n \geq 1\}$ and, hence, $\sigma(L) = \{n \cdot \ell_j^{-1} : n \geq 1, j \geq 1\}$. Therefore, the spectral zeta function of $L$ (or, equivalently, of the Dirichlet Laplacian on $\Omega$), denoted by $\zeta$, and defined by $\zeta(s) = \sum_{f \in \sigma(L)} f^{-s}$, for large $\text{Re}(s)$, is given by the following product formula:

$$\zeta(s) = \zeta(s) \cdot \zeta_L(s),$$

(2.1)

where $\zeta_L$ is the geometric zeta function of $L$ (given by its expression in §2a or by its meromorphic continuation thereof). Formula (2.1) was first observed in [14,15]. It, along with its counterpart, was key to our later developments. (We note that formula (2.1) was later reinterpreted and extended to various settings, in terms of complex dynamics; see [34–36].) At the level of the counting functions, (2.1) becomes $N_v(x) = \sum_{n=1}^\infty N_L(x/n)$, for $x > 0$, where $N_L$ (resp., $N_v$) denotes the geometric (resp., spectral or frequency) counting function of $L$, counting, respectively, the number of reciprocal lengths $\ell_j^{-1}$ or of frequencies $f$ not exceeding $x > 0$: $N_L(x) = \sum_{\ell_j^{-1} \leq x} 1$ and $N_v(x) = \sum_{f \in \sigma(L)} 1$, where both the lengths and the frequencies (essentially, the square roots of the eigenvalues) are counted according to their multiplicities [7, theorem 1.21].

We state some of the basic properties of $\zeta = \zeta(s)$ used throughout this paper, most often implicitly. The Riemann zeta function has a meromorphic extension to all $\mathbb{C}$, denoted by $\zeta = \zeta(s)$.

Furthermore, for all $s \in \mathbb{C}$ such that $\text{Re}(s) > 1$, $\zeta(s)$ is given by the classic Dirichlet series $\sum_{n=1}^\infty n^{-s}$ (used to prove equation (2.1)) and Euler product $\prod_{p \in \mathbb{P}} (1 - p^{-s})^{-1}$, where $\mathbb{P}$ denotes the set of prime numbers. It follows that $\zeta(s) \neq 0$ for $\text{Re}(s) > 1$. Therefore, in light of the functional equation satisfied by $\zeta$ (namely, $\zeta(s) = \z(1 - s)$, where $\z(s) := \pi^{-s/2} \Gamma(s/2)\z(s)$ denotes the completed or global Riemann zeta function and $\Gamma$ is the classic gamma function), it follows that $\zeta(s) \neq 0$ for $\text{Re}(s) < 0$, except at the trivial zeros of $\zeta$, located at $s = -2, -4, -6, \ldots$. Also, according to Hadamard’s theorem (and the functional equation), we have $\zeta(s) \neq 0$ for $\text{Re}(s) = 1$ (and hence also for $\text{Re}(s) = 0$). Therefore, the other zeros of $\zeta$ (called the critical or non-trivial zeros) must lie in the (open) critical strip $0 < \text{Re}(s) < 1$ (e.g. [33,37–40] for these and other key properties of $\zeta$).

(d) Complex dimensions of fractal strings, via explicit formulae

Let $L = \{\ell_j\}_{j=1}^\infty$ be a given fractal string and $\zeta_L = \zeta_L(s)$ denote its geometric zeta function, given in §2a for $\text{Re}(s) > D$, where $D = D_L$ is the Minkowski dimension of $L$. Briefly, the complex dimensions of $L$ (or of $\partial \Omega$, where $\Omega$ is any geometric representation of $L$) are the poles of the meromorphic continuation of $\zeta_L$. More precisely, let $U \subset \mathbb{C}$ be a domain to which $\zeta_L$ can be meromorphically continued. Then, the visible complex dimensions of $\zeta_L$ (in $U$) are the poles of $\zeta_L$ in $U$. If $U = \mathbb{C}$ or when no ambiguity may arise, we just call them the complex dimensions of $L$. We denote by $D = D_L$ the set of complex dimensions of $L$. In light of the definition of the abscissa of convergence of $\zeta_L$ (§2a), $\zeta_L$ is holomorphic for $\text{Re}(s) > D$. In fact, it can be shown (using well-known results about Dirichlet series, see [22, §VI.2.3]) that $\zeta_L(s) \to +\infty$ as $s \to D^+$, $s \in \mathbb{R}$. Therefore, $\{\text{Re}(s) > D\}$ is the largest open right half-plane (of the form $\{\text{Re}(s) > \alpha\}$, for some $\alpha \in \mathbb{R} \cup \{\pm \infty\}$ to which $\zeta_L$ can be holomorphically continued. (Furthermore, $\{\text{Re}(s) > D\}$ is the largest such half-plane in which the series $\sum_{j=1}^\infty \ell_j^s$ is convergent or, equivalently, absolutely convergent; e.g. [22, §VI.3].) It follows, in particular, that we must always have $D \subseteq \{\text{Re}(s) \leq D\}$. If, in addition, $\zeta_L$ can be meromorphically extended to an open neighbourhood of $D$, then $D \in \mathbb{D}$ and $D = \max\{\text{Re}(s) : \omega \in \mathbb{D}\}$. Here and thereafter, we (often) use the short-hand notation $\{\text{Re}(s) > \alpha\} := \{s \in \mathbb{C} : \text{Re}(s) > \alpha\}$. Similarly, if $\alpha \in \mathbb{R}$, $\{\text{Re}(s) = \alpha\}$ stands for the vertical line $\{s \in \mathbb{C} : \text{Re}(s) = \alpha\}$ and analogously for the vertical strip $[\alpha \leq \text{Re}(s) \leq \beta]$, with $\alpha \leq \beta$. 
Our aim here is not to provide a complete description of the theory of complex dimensions [5–7] (and especially [7]). Instead, we just provide a few examples and results. In the present exposition, we reverse the chronological order as the results from [1–4] (§3) were obtained before the rigorous theory of complex dimensions [5–7], and in fact (along with the work in [13–15,31], in particular) provide a natural heuristic and mathematical motivation for it.

Let \( \mathcal{L} = \text{CS} \) be the Cantor string; it is associated with the bounded open set \( \Omega \subset \mathbb{R} \) defined as the complement in \([0, 1]\) of the classic (ternary) Cantor set \( C \). Note that \( \Omega \) is merely the union of the sequence of disjoint open intervals (the ‘middle thirds’) which are deleted in the standard construction of \( C \). So that \( \mathcal{L} \) is given by the sequence of lengths \( \frac{1}{3}, \frac{1}{9}, \frac{1}{27}, \frac{1}{81}, \cdots \) repeated with multiplicity \( 2^n-1 \) for \( n = 1, 2, 3, \ldots \). Hence, \( \zeta \) has a meromorphic continuation to all \( \mathbb{C} \) given by \( \zeta(s) = 3^{-s} (1 - 2^{-s}) \), for all \( s \in \mathbb{C} \). It follows that the set of complex dimensions of \( C \) (or, equivalently, of the Cantor set \( C = \partial \Omega \)) is given by \( \mathcal{D}_{\text{CS}} = \{ D + i n \mid n \in \mathbb{Z} \} \), where \( i := \sqrt{-1}, \ D = \log_3 2 \) is the Minkowski dimension of \( \text{CS} \) (or, equivalently, of \( C \)) and \( p := 2\pi / \log_3 \) is the oscillatory period of \( \text{CS} \). Here and in the sequel, we let \( \mathbb{Z} := \{ 0, \pm 1, \pm 2, \ldots \} \).

In [7, ch. 5, 6 and 8], general explicit formulae expressing, for instance, the geometric and spectral counting functions \( N_{\mathcal{L}} = N_{\mathcal{L}}(x) \) and \( N_{\nu} = N_{\nu}(x) \), respectively, as well as the volume (i.e. length) \( V = V(\epsilon) := |\{ x \in \Omega : \text{dist}(x, \partial \Omega) < \epsilon \}| \) of the (inner) \( \epsilon \)-tubes of a fractal string \( \mathcal{L} \), in terms of the underlying complex dimensions of \( \mathcal{L} \) are given. For example, under suitable growth conditions on \( \zeta(L) \), the general fractal tube formula obtained in [7, ch. 8] is of the following form (for notational simplicity, we assume here that all the poles are simple):

\[
V(\epsilon) = \sum_{\omega \in \mathcal{D}_{\mathcal{L}}} \text{res}(\zeta_L; \omega) \frac{(2\epsilon)^{1-\omega}}{\omega(1-\omega)} + R(\epsilon), \tag{2.2}
\]

where the error term \( R = R(\epsilon) \) can be explicitly estimated. Here and in the sequel, \( \text{res}(\zeta_L; \omega) \) denotes the residue of the meromorphic function \( \zeta_L = \zeta_L(s) \) evaluated at \( s = \omega \).

For a self-similar string, we have \( R(\epsilon) \equiv 0 \) in (2.2) and so the formula is exact. In particular, for the above Cantor string \( \text{CS} \), we have \( e^{-1-D} V_{\text{CS}} = G(\log_3 e^{-1}) - 2e^D \), where \( G \) is a non-constant periodic function (of period 1) which is bounded away from 0 and infinity. Therefore, the Cantor string \( \text{CS} \) (and, hence, also the Cantor set \( C \)) is not Minkowski measurable (as shown in [2] via a direct computation) and its lower and upper Minkowski contents are given by \( \mathcal{M}_- = \min G > 0 \) and \( \mathcal{M}^* = \max G < \infty \), respectively (with explicit values first calculated in [2], see also [7, eqn (1.12)]).

At the level of the counting functions \( N_{\mathcal{L}} \) and \( N_{\nu} = N_{\nu, CS} \), we have the following explicit formulae, still for the Cantor string \( \text{CS} [7, \S 1.2.2 \text{ and } \S 6.4.3] \):

\[
N_{\mathcal{L}}(x) = \frac{1}{2 \log 3} \sum_{n \in \mathbb{Z}} \frac{x^{D+i n \log_3}}{D+in} - 1
\]

and

\[
N_{\nu, \text{CS}}(x) = x + \frac{1}{2 \log 3} \sum_{n \in \mathbb{Z}} \zeta(D + i n \log_3) x^{D+i n \log_3}. 
\]

The first explicit formula was obtained by Riemann [32] and is certainly one of the most beautiful formulae in mathematics. It expresses the counting function of the primes in terms of the zeros and the (single) pole of \( \zeta \) (i.e. in terms of the poles of \( -\zeta'/\zeta \)). A modern version of Riemann’s explicit formula can be stated as follows: let \( N(x) := \sum_{\rho \leq x} 1/n \) be the function counting all the prime powers \( p^n \leq x \) with a weight \( 1/n \). Then

\[
N(x) = \text{Li}(x) - \sum_{\rho} \text{Li}(x^{\rho}) + \int_{x}^{+\infty} 1 \frac{dx}{x^2 - 1} \log x - \log 2,
\]

where the sum is taken over the zeros \( \rho \) of \( \zeta \) and \( \text{Li}(x) := \int_0^x (dt/\log t) \) is the integral logarithm (e.g. ([7, pp. 3 and 141], [33, \S 1.16], [37,39]) or [40]). The explicit formulae obtained in [7] extend this explicit formula to a general setting where the corresponding zeta function does not typically satisfy a functional equation or have an Euler product representation [7, ch. 5–11].
From our point of view, an important application of the general explicit formulae of [7] is that if we write \( \eta_L = \sum_{j=1}^{\infty} \delta_{\xi_j} \) and \( \nu_L = \sum_{\omega \in D_L} \delta_{\omega} \), where \( \delta_{\omega} \) denotes the Dirac measure (or distribution) at \( \omega \), then (under suitable assumptions and if the complex dimensions are simple, see [7, §6.3.1]) we have the distributional identities: 
\[
\eta_L = \sum_{\omega \in D_L} \text{res}(\zeta_L; \omega) x^{\omega - 1} \quad \text{and} \quad \nu_L = \zeta(1) + \sum_{\omega \in D_L} \text{res}(\zeta_L; \omega) x^{\omega - 1},
\]
where we have neglected the error term. In [7, §6.3.1], these formulae are, respectively, called the density of geometric states and the density of spectral states. At the level of the measures, the spectral operator, which will play a key role in the sequel (see, especially, §§4 and 5), can be intuitively interpreted (as in [7, §6.3.2]) as the operator sending the geometry (represented by \( \eta_L \)) onto the spectrum (represented by \( \nu_L \)) of a fractal string \( L \). In light of the above explicit formulae for \( \eta_L \) and \( \nu_L \), it amounts (essentially) to multiplying by \( \zeta(\omega) \) the local term \( \text{res}(\zeta_L; \omega) x^{\omega - 1} \) associated with each complex dimension \( \omega \in D_L \). In §4a, we will discuss another heuristic interpretation of the spectral operator, based on [7, §6.3.2] and given at the level of the counting functions \( N_L \) and \( N_{\nu} \), while in §4b we will give a rigorous definition of the spectral operator (obtained in [8–12]).

Remark 2.1.

(a) At a fundamental level, the theory of complex dimensions developed in [5–7] is a theory of geometric, spectral, dynamical or arithmetic oscillations (or ‘vibrations’). More concretely, as is apparent in the explicit formulae discussed in this subsection (see, for example, equation (2.2)), the real (resp., imaginary) parts of the complex dimensions are associated with the amplitudes (resp., frequencies) of these oscillations (viewed at the geometric level, for example, as waves propagating in the space of ‘scales’).

(b) In [5–7], an object is said to be fractal if its associated zeta function has at least one non-real complex dimension (i.e. pole), with positive real part. Accordingly, all self-similar fractal strings are fractals. Furthermore, the Cantor curve (or ‘devil’s staircase’) is fractal in this new sense, whereas (contrary to everyone’s geometric intuition) it is not fractal in the sense of Mandelbrot’s definition [41] (which according to the Hausdorff dimension must be strictly greater than the topological dimension). We refer to the definition of fractality provided in [7, §12.1 and §12.2]. We note that, in [7, §13.4.3], the above definition of fractality is extended to allow for the zeta function to have a natural boundary along some suitable curve (i.e. not to have a meromorphic continuation beyond that curve); such objects are now called hyperfractals in [42–48]. Recently, in [42–45], the authors have constructed maximally hyperfractal strings (and also compact subsets of \( \mathbb{R}^N \), for any \( N \geq 1 \), in the sense that the geometric (or fractal) zeta functions have singularities at every point of the vertical line \( \text{Re}(s) = D \).

(c) The mathematical theory of complex dimensions of fractal strings has many applications to a variety of subjects, including fractal geometry, spectral geometry, number theory, arithmetic geometry, geometric measure theory, probability theory, dynamical systems and mathematical physics (e.g. [1–15,20,27–31,34–36,42–74] and references therein). We note that recent developments in the theory are described in [7, ch. 13], including a first attempt at a higher dimensional theory of complex dimensions for the special case of fractal sprays (in the sense of [30]) and self-similar tilings (see [7, §13.1], based on [63–66,72]), \( p \)-adic fractal strings and associated non-Archimedean fractal tube formulae (see [7, §13.2], based on [56–60]), multi-fractal zeta functions and their ‘tapestries’ of complex dimensions (see [7, §13.3], based on [50,55,67]), random fractal strings (such as stochastically self-similar strings and the zero-set of Brownian motion) and their spectra (see [7, §13.4], based on [51]), as well as a new approach to the RH based on a conjectural Riemann flow of fractal membranes (i.e. quantized fractal strings) and correspondingly flows of zeta functions (or ‘partition functions’) and of the associated zeros (see [7, §13.5], which gives a brief overview of the aforementioned book [20], In search of the Riemann zeros).

(d) Recently, in [42–48], the theory of complex dimensions of fractal strings developed in [5–7] has been extended to any bounded subset of Euclidean space \( \mathbb{R}^N \) (and even to
3. Direct and inverse spectral problems for fractal strings

In this section, we study direct (§3a) and inverse (§3b) spectral problems for fractal strings, based on the work in [1,2] and [3,4], respectively, and explore in the one-dimensional case some of the ramifications of the work in [13] on a partial resolution of the Weyl–Berry conjecture for fractal drums (see the discussion surrounding equation (3.1)); see also [27] for a recent and more detailed survey on this topic. In the process, we will be led naturally to establishing close connections between these problems and the values of the Riemann zeta function $\zeta = \zeta(s)$ in the critical interval $0 < s < 1$ (see §3a) as well as with the critical zeros (see §3b); that is, the zeros of $\zeta(s)$ located in the critical strip $0 < \Re(s) < 1$.

(a) Minkowski measurability of fractal strings and spectral asymptotics with a monotonic second term

In [2] (announced in [1]), the author and Carl Pomerance have established (in the case of fractal strings, that is, in dimension one) the modified Weyl–Berry conjecture (MWB conjecture) formulated by the author in [13] and according to which if a fractal string $L = \{\ell_j\}_{j=1}^\infty$ of Minkowski dimension $D \in (0, 1)$ is Minkowski measurable with Minkowski content $M$ (see the discussion at the end of §2a), then its spectral (or frequency) counting function $N_\nu = N_{\nu,L}$ admits a monotonic asymptotic second term, of the form $-c_D M x^D$, where $c_D$ is a positive constant depending only on $D$:

$$N_\nu(x) = W(x) - c_D M x^D + o(x^D), \quad \text{as } x \to +\infty,$$  \tag{3.1}

where the Weyl (or leading) term $W$ is given by $W(x) := |\Omega|x$, with $|\Omega| = \sum_{j=1}^\infty \ell_j$ being the ‘volume’ (or, really, the ‘length’) of any geometric realization $\Omega$ of $L$.

Theorem 3.1 (Resolution of the MWB conjecture for fractal strings, [2]). If $L$ is a Minkowski measurable fractal string of dimension $D \in (0, 1)$, then (3.1) holds, with the Weyl term given by $W(x) := |\Omega|x$ as above and the positive constant $c_D$ given by $c_D := (1 - D)2^{D-1}(-\zeta(D))$.

Note that $c_D > 0$ because the Riemann zeta function is negative in the critical interval $(0, 1)$ (e.g. [40]): $\zeta(D) < 0$ as $D \in (0, 1)$. The proof of theorem 3.1 provided in [2] is given in two different steps. The first step consists in proving the following key characterization of Minkowski measurability for fractal strings (i.e. in one dimension), due to [2] and also of independent geometric interest:

The fractal string $L = \{\ell_j\}_{j=1}^\infty$ (of Minkowski dimension $D \in (0, 1)$) is Minkowski measurable if and only if $\ell_j \sim L^{-1/D}$ as $j \to \infty$, for some $L \in (0, +\infty)$; i.e. iff $L := \lim_{j \to \infty} \ell_j \cdot j^{1/D}$ exists in $(0, +\infty)$. In that case, its Minkowski content $M$ is given by $M = 2^{1-D}L/(1 - D)$.

The second step of the proof of theorem 3.1, also due to [2], consists in first observing that (as the intervals or substrings of which the fractal string $\Omega$ is composed are vibrating independently of one another) $N_\nu(x) = \sum_{j=1}^\infty [\ell_j x]$, where $[y]$ is the integer part of $y \in \mathbb{R}$, and then establishing that the following (monotonic) two-term asymptotic expansion holds: if $\ell_j \sim L^{-1/D}$ as $j \to \infty$, then $\sum_{j=1}^\infty [\ell_j x] = (\sum_{j=1}^\infty \ell_j) x + \zeta(D) x^D + o(x^D)$, as $x \to +\infty$. (Observe that the leading term in this formula coincides with the Weyl term, $W(x) = |\Omega|x$.) In the proof given in [1,2], one uses, in particular, the following form of the analytic continuation of $\zeta = \zeta(s)$ to the open right half-plane $\Re(s) > 0$ (e.g. [22, VI.3.2]): $\zeta(s) = 1/(s - 1) + \int_0^{+\infty} (t^{-s} - t^{-s}) \, dt$, for $\Re(s) > 0$. Note that, on the right-hand side of this identity, the first term yields the simple, simple pole of $\zeta$ at $s = 1$, while the second term defines a holomorphic function of $s$ for $\Re(s) > 0$. This identity also plays an important role in [8,12]; see §4c. Proofs of a part of step 1 above were provided in [52,76], using

‘relative fractal drums’ of $\mathbb{R}^N$ for any $N \geq 1$, via the use of new ‘fractal zeta functions’; namely, the ‘distance zeta function’ (introduced by the author in 2009) and the closely related ‘tube zeta function’. For a comprehensive exposition of the resulting theory, we refer to the forthcoming book [42], along with the papers [43–47] and survey articles [48,75].
different techniques from the theory of dynamical systems or from geometric measure theory, respectively.

Remark 3.2.

(a) The original Weyl conjecture was stated (in 1912) for ‘nice’ smooth (or piecewise smooth) bounded domains $\Omega \subset \mathbb{R}^N (N \geq 1)$ [77], where the mathematician Hermann Weyl also obtained his classic formula (see, for example, [78,79] or [80, vol. IV]) for the leading spectral asymptotics of the Dirichlet (or Neumann) Laplacian on $\Omega$ in the case of a (piecewise) smooth boundary $\partial \Omega$. The original and stimulating Berry conjecture was stated (in the late 1970s) by the physicist Michael Berry [81,82] for bounded open sets $\Omega \subset \mathbb{R}^N$ with ‘fractal’ boundary. It was expressed in terms of the Hausdorff dimension and measure of the boundary, and that form was disproved in [83]; see also [13], for general mathematical reasons and with a simple family of counterexamples, valid in any dimension. A partial resolution of the Weyl–Berry conjecture for fractal drums was obtained by the author in [13], where it was proved that for the Dirichlet (as well as, under suitable hypotheses, the Neumann) Laplacian on a bounded open set $\Omega \subset \mathbb{R}^N$ (for $N \geq 1$) with boundary $\partial \Omega$ satisfying $D = D(\partial \Omega) \in (N - 1, N)$ and $\mathcal{M}^* < \infty$, we have

$$N_\nu(x) = W(x) + O(x^D), \quad \text{as } x \to +\infty,$$

(3.2)

where $W(x) := C_N |\Omega|_{N \times N}$ is the $N$-dimensional Weyl term and $|\Omega|_N$ is the $N$-dimensional volume of $\Omega$. Here, $D$ notes the (upper) Minkowski dimension of $\partial \Omega$ and, as noted in [13], we always have $D \in [N - 1, N]$. The error estimate in (3.2) was also shown to be sharp for every $N \geq 1$ and every $D \in (N - 1, N)$; see [13, examples 5.1 and 5.1']. In the non-fractal case when $D = N - 1$, the error term in the counterpart of (3.2) involves a logarithm term; see [13] and the earlier work by Métivier [84]. For further results concerning the (modified) Weyl–Berry conjecture and discussion of its many physical applications [81,82], see ([7, §12.5 and appendix B], [13–15], [27, §4.1], [42,48]) and references therein, including [1–4,30,31,51,83,85–89]. (We note that the analogue of the MWB conjecture is not true, in general, when $N \geq 2$, although no counterexample seems to be known for simply connected planar domains; see [85] and, especially, [30].) For the case of ‘drums with fractal membrane’ (Laplacians on fractals [90], rather than on open sets with fractal boundary [13–15]), see [15,91] along with, for example, [34–36,49,92–96] and references therein.

(b) It is shown in [2] that a fractal string (of dimension $D \in (0, 1)$) is Minkowski non-degenerate (i.e. $0 < \mathcal{M}_*(\leq) \mathcal{M}^* < \infty$) if and only if $\ell_j \approx j^{-1/D}$ as $j \to \infty$ and $\psi(x) := |N_\nu(x) - W(x)| \approx x^D$ as $x \to +\infty$, where the symbol $\approx$ means that we have two-sided error estimates; e.g. $a x^D \leq \psi(x) \leq b x^D$, for some constants $a, b > 0$. In §3b, the situation is very different when ‘Minkowski non-degeneracy’ is replaced by ‘Minkowski measurability’.

(b) The sound of fractal strings and the Riemann zeros

In §3a, we have studied a direct spectral problem for fractal strings as in theorem 3.1 based on [2] (as well as in the corresponding MWB conjecture from [13]), we assume something about the geometry of the fractal string $L$ (namely, the Minkowski measurability of $L$) and deduce some information about the spectrum of $L$ (namely, that $N_\nu(x)$ admits a monotonic asymptotic second term). Conversely, it is natural to consider (as in [3,4]) the following associated inverse spectral problem:

(ISP)$_D$ If the fractal string (of Minkowski dimension $D \in (0, 1)$) is such that its spectral counting function $N_\nu(x)$ admits a monotonic asymptotic second term as $x \to +\infty$, proportional to $x^D$ (i.e. of the form $C x^D$, where the non-zero constant $C$ depends only on $D$ and $L$, compare with (3.1)), is it true that $L$ is necessarily Minkowski measurable?
As was shown by the author and Helmut Maier in [4] (announced in [3]), it turns out that this question (à la Mark Kac, but of a very different nature from that in [79]) ‘Can one hear the shape of a fractal string?’ is intimately related to the existence of critical zeros of $\zeta = \zeta(s)$ on the vertical line $(\text{Re}(s) = D)$, and therefore to the RH. In the sequel, we will say that the inverse spectral problem (ISP)$_D$ has an affirmative answer for a given value of $D \in (0,1)$ if the above fractal strings version of Kac’s question has an affirmative answer for all fractal strings of dimension $D$. Equivalently, (ISP)$_D$ has a negative answer for some given $D \in (0,1)$ if there exists a non-Minkowski measurable fractal string $\mathcal{L}$ of (Minkowski) dimension $D$ such that its spectral counting function $N_\nu = N_{\nu,\mathcal{L}}$ has a monotonic asymptotic second term (proportional to $x^D$). Note, however, that, in light of the result of [2] discussed in remark 3.2(b), such a fractal string $\mathcal{L}$ must be Minkowski non-degenerate.

Theorem 3.3 (Riemann zeros and the inverse problem (ISP)$_D$, [4]). Fix $D \in (0,1)$. Then, the inverse spectral problem (ISP)$_D$ has an affirmative answer for this value of $D$ if and only if $\zeta = \zeta(s)$ does not have any zeros on the vertical line $(\text{Re}(s) = D)$.

It follows from theorem 3.3 that, in the midfractal case when $D = \frac{1}{2}$, the inverse problem (ISP)$_D$ has a negative answer as $\zeta$ has a zero on the critical line $(\text{Re}(s) = \frac{1}{2})$. (Actually, $\zeta(s)$ vanishes infinitely often for $\text{Re}(s) = \frac{1}{2}$ but this result, due to Hardy (see [40]), is not needed here. One zero, and its complex conjugate, which is also a zero of $\zeta$ since $\zeta(\bar{s}) = \bar{\zeta}(s)$, suffices.) Furthermore, any counterexample to the RH would show that (ISP)$_D$ fails to have an affirmative answer for some $D \neq \frac{1}{2}$. (None is expected, however.) Hence, we obtain the following result, which is really a corollary of theorem 3.3 (in light of the functional equation satisfied by $\zeta = \zeta(s)$, which exchanges $s$ and $1 - s$) but that we state as a theorem because it is the central result of [3,4] and is a key motivation for several of the main results stated in §§4 and 5.

Theorem 3.4 (RH and inverse spectral problems for fractal strings, [4]). The inverse spectral problem (ISP)$_D$ has an affirmative answer for all $D \in (0,1)$ other than $D = \frac{1}{2}$ (or, equivalently, for all $D \in (0, \frac{1}{2})$ or, else, for all $D \in (\frac{1}{2}, 1)$, respectively) if and only if the RH is true.

Theorem 3.4 provides a spectral and geometric reformulation of the RH. For the purpose of this paper, we will refer to theorem 3.4 (as well as to theorem 4.8, a corresponding operator theoretic version of theorem 3.4 obtained in [8–10] and discussed in §4d) as a symmetric criterion for the RH. Obtaining an asymmetric criterion for RH will be the objective in §5.

Remark 3.5. The proof of theorem 3.4 relies on the Wiener–Ikehara Tauberian theorem [97] (for one of the implications) and (for the reverse implication) on the intuition of the notion of complex dimension which, at the time, was still conjectural [1,2,14,15]. On the other hand, theorem 3.4, along with results and conjectures in [1,2,13–15,30,31,91], in particular, provided some of the key motivations for developing [5–7] a rigorous theory of complex (fractal) dimensions via (generalized) explicit formulae and geometric zeta functions. In [5–7], especially [7, ch. 9], theorem 3.4 is reformulated in terms of explicit formulae and extended to all arithmetic zeta functions for which the GRH is expected to hold. For information about GRH, see for example, ([98,99], [20, appendices. B, C and E]) or [7, appendix A].

4. The spectral operator as a quantized Riemann zeta function

In this section, we give a brief overview of the work of Hafedh Herichi and the author on the aspects of quantized number theory. This work is presented in a series of research and survey articles [9–12], and described in detail in [8]. More specifically, in [8–12], the main object of investigation is the spectral operator, sending the geometry onto the spectrum of a fractal string. Originally seen as acting on the space of (generalized) fractal strings, it is now viewed mathematically (and physically) as a suitable quantization of the Riemann zeta function: $\alpha = \zeta(\partial)$, where $\partial$ is a suitable version of the differentiation operator $d/dt$, called the infinitesimal shift of the real line and acting on the Hilbert space $\mathbb{H}_c = L^2(\mathbb{R}, e^{-2\pi t} dt)$, a weighted $L^2$-space; see §4b. Accordingly, the spectral operator $\alpha = \alpha_c$ depends on a parameter $c$ which can be thought of...
heuristically as providing an upper bound for the Minkowski dimensions of the fractal strings on which the operator acts. In fact, as we shall see in §4a, it is more convenient to replace the fractal strings themselves by their associated (geometric and spectral) counting functions. As it turns out, the spectral operator, \( a = \zeta(\partial) \), which was originally introduced at the semi-heuristic level in [7, §6.3.1 and §6.3.2] (see also [7, §6.3.1 and §6.3.2]) satisfies (at the operator theoretic level) most of the properties of the classic Riemann zeta function \( \zeta = \zeta(s) \), including a quantized Dirichlet series, a quantized Euler product and an operator-valued ‘analytic continuation’. (See §4c, along with ([8, ch. 7], [12]).) One of the key motivations of the work in [8–12] is to obtain a rigorous functional analytic version of the work of the author and H. Maier in [4] on inverse spectral problems for fractal strings and the RH briefly described in §3b. This goal is achieved (see §4d) by studying the quasi-invertibility of the spectral operator as a function of the parameter \( c \) in the critical interval \((0, 1)\), which is the natural range of possible dimensions of fractal strings, leaving aside the least and most fractal cases. In this light, the RH is true if and only if the midfractal case \( c = \frac{1}{2} \) is the only exception to the quasi-invertibility of \( a \); see §4d.

(a) Heuristic spectral operator

As we have seen in §2c, the analogue (at the level of the counting functions) of the factorization formula (2.1), connecting the geometric and the spectral zeta functions of a fractal string \( \mathcal{L} \), via the Riemann zeta function \( \zeta \), is provided by the following formula, connecting the geometric and spectral counting functions \( N_\mathcal{L} \) and \( N_\nu \), respectively, of a given fractal string \( \mathcal{L} \); \( N_\nu(x) = \sum_{n=1}^\infty N_\mathcal{L}(x/n) \), for all \( x > 0 \). (Note that, for a fixed \( x > 0 \), only finitely many terms contribute to this sum; however, the number of these terms tends to infinity as \( x \to +\infty \).) The (heuristic) spectral operator, at the level of the counting functions, is then given by the map \( g(x) = N_\mathcal{L}(x) \mapsto N_\nu(g)(x) = \sum_{n=1}^\infty g(x/n) \). Hence, it can be thought of as sending the geometry onto the spectrum of a fractal string \( \mathcal{L} \). At an even more fundamental level, and using the notation of §2d, the heuristic spectral operator can be thought of as being the map \( \eta := \sum_{j=1}^\infty \delta_{\zeta_j^{-1}} \mapsto \nu := \sum_{\zeta \in \mathcal{L}} \delta_\zeta \), where \( \sigma(\mathcal{L}) \) denotes the (frequency) spectrum of \( \mathcal{L} \) (given at the beginning of §2c) and the measures \( \eta \) and \( \nu \), respectively, represent the geometry and the spectrum of \( \mathcal{L} \), viewed as generalized fractal strings (in the sense of [7, ch. 4]). (For a brief discussion of the spectral operator expressed in terms of the complex dimensions of the fractal string \( \mathcal{L} \), see the text immediately preceding remark 2.1 about the densities of geometric and of spectral states of \( \mathcal{L} \); see also [6, §6.3.1].)

The spectral operator was first introduced (in this semi-heuristic context) in [6, §6.3.2] (see also [7, §6.3.2]). The above version is referred to in [8–12] as the multiplicative version of the (heuristic) spectral operator. From now on, we will only work with the (equivalent) additive version, which we next describe.

Viewed additively (i.e. after having made the change of variable \( x = e^t, t = \log x \), with \( x > 0 \) and \( t \in \mathbb{R} \), we obtain the additive spectral operator \( a \), to be briefly referred to henceforth as the (heuristic) ‘spectral operator’:

\[
f(t) \mapsto a(f)(t) = \sum_{n=1}^\infty f(t - \log n),
\]

(4.1)

where the functions \( f = f(t) \) are viewed as functions of the variable \( t \in \mathbb{R} = (-\infty, +\infty) \). (The precise mathematical setting will be specified in §4b.) Given a prime \( p \) (i.e. \( p \in \mathcal{P} \), where \( \mathcal{P} \) denotes the set of all prime numbers), the associated local (operator-valued) Euler factor \( a_p \) is given by \( f(t) \mapsto \sum_{m=0}^\infty f(t - m \log p) \). Formally, the spectral operator \( a \) is connected to its local Euler factors \( a_p \) (with \( p \in \mathcal{P} \)) via the following Euler product representation: \( a = \prod_{p \in \mathcal{P}} a_p \), where the infinite product is to be understood in the sense of the composition of operators. Next, assuming for now that the function \( f \) is infinitely differentiable, we formally define the differentiation operator \( \partial = d/dt \) (so that \( \partial f = f' \), the derivative of \( f \), \ldots, \( \partial^k f = f^{(k)} \), the \( k \)th derivative of \( f \), for any integer \( k \geq 0 \)). Then, \( e^{-\hbar \partial} \) acts as a shift or translation operator. Namely, \( f(t - h) = (e^{-\hbar \partial})(f)(t) \).
We conclude this discussion by giving the heuristic motivations for the above Euler product formula and for the following Dirichlet series representation of \( a \), which will serve as the basis for our rigorous discussion in §4b,c: \( a = \sum_{n=1}^{\infty} (-1)^{n-\delta} = \zeta(\delta) \). In light of the above discussion, we have for any prime \( p \in \mathbb{P} \), \( a_p = \sum_{m=0}^{\infty} e^{-m(\log p)^2} = \sum_{m=0}^{\infty} (p^{-\delta})^m = (1 - p^{-\delta})^{-1} \) and so \( a = \sum_{n=1}^{\infty} e^{-(\log n)^2} = \sum_{n=1}^{\infty} (n^{-\delta}) = \zeta(\delta) \), as desired. All these formulae should be verified by applying the corresponding expression to the function \( f \) and evaluating at a generic \( t \in \mathbb{R} \). In actuality, the rigorous justification of these heuristic formulae requires a lot more work and is provided (under suitable assumptions) in [8–12], as we next briefly discuss in §4b,c.

(b) The infinitesimal shift \( \partial \) and the spectral operator \( a = \zeta(\partial) \)

Given \( c \in \mathbb{R} \), let \( \mathbb{H}_c = L^2(\mathbb{R}, e^{-2ct} \, dt) \) be the complex Hilbert space of (complex-valued, Lebesgue measurable) square-integrable functions with respect to the absolutely continuous measure \( \mu(dt) = e^{-2ct} \, dt \). The inner product of \( \mathbb{H}_c \) is given by \( \langle f, g \rangle_c := \int_{-\infty}^{\infty} e^{-2ct} f(t) \overline{g(t)} \, dt \) and, thus, the associated norm \( \|f\|_c \) is naturally given by \( \|f\|_c^2 = \langle f, f \rangle_c = \int_{-\infty}^{\infty} e^{-2ct} |f(t)|^2 \, dt < \infty \).

Let \( C_{\text{abs}}(\mathbb{R}) \) denote the space of (locally) absolutely continuous functions on \( \mathbb{R} \); e.g. [100] or [101]. Let \( \partial = \partial_c = d/dt \) denote the differentiation operator, acting on \( \mathbb{H}_c \) and with domain \( D(\partial) = \{ f \in \mathbb{H}_c \cap C_{\text{abs}}(\mathbb{R}) : f' \in \mathbb{H}_c \} \), where \( f' \) denotes the distributional (or weak) derivative of \( f \) (e.g. [102–105]), which (since \( f \) is absolutely continuous) can be interpreted as the usual pointwise derivative of \( f \) existing (Lebesgue) almost everywhere (a.e., in short) on \( \mathbb{R} \). Then, for every \( f \in D(\partial) \), we have \( \partial f := f' \). The operator \( \partial \) so defined is called the infinitesimal shift of the real line. (Note that it depends on the choice of the parameters \( c \in \mathbb{R} \).)

**Theorem 4.1 ([8]).** For every \( c \in \mathbb{R} \), the infinitesimal shift \( \partial = \partial_c \) is an unbounded normal operator on \( \mathbb{H}_c \), with spectrum \( \sigma(\partial) \) given by \( \sigma(\partial) = \{ \Re(s) = c \} = c + i \mathbb{R} \), the vertical line of abscissa \( c \). (Here, \( i := \sqrt{-1} \).) Furthermore, its adjoint \( \partial^* \) is given by \( \partial^* = 2c - \partial \); so that we also have \( \sigma(\partial^*) = \{ \Re(s) = c \}. \)

In addition, neither \( \partial \) nor \( \partial^* \) has any eigenvalues (hence, its point spectrum is empty).

Recall that given a (closed) unbounded, linear operator \( L \) on a complex Hilbert space \( H \), its spectrum, denoted by \( \sigma(L) \), is defined as the set of all \( \lambda \in \mathbb{C} \) such that \( L - \lambda I \) is not invertible, where \( I \) denotes the identity operator. (In finite dimensions, this amounts to requiring that \( L - \lambda I \) is not 1–1, and, hence, \( \lambda \) is an eigenvalue of \( L \); so that \( \sigma(L) \) is the finite set of eigenvalues of \( L \) in that case.) The spectrum \( \sigma(L) \) is always a closed subset of \( \mathbb{C} \). Furthermore, \( L \) is invertible if and only if \( 0 \not\in \sigma(L) \). Moreover, an (unbounded, linear) operator \( M \), with domain \( D(M) \), is said to be invertible if the linear map \( M : D(M) \to H \) is bijective and its set theoretic inverse \( M^{-1} : H \to D(M) \subseteq H \) is bounded (which, by the closed graph theorem, is automatically true if \( M \) is closed, as will always be the case in this paper). Finally, if \( M \) is closed (i.e. if its graph is a closed subspace of \( H \times H \) for the graph norm), then it is said to be normal if \( M^* M = M M^* \) (equality between unbounded operators), where \( M^* \) denotes the adjoint of \( M \). For all of these notions, see ([80, vol. I], [103,104,106,107]).

In the next result, which is needed (in particular) to rigorously justify the formulae provided in §4a (see §4c), we will implicitly use the theory of strongly continuous contraction semigroups of bounded linear operators and their infinitesimal generators (i.e. \( m \)-accretive operators; e.g. [80, 103,107–109]).

**Proposition 4.2 ([8]).** The semigroup \( \{ e^{-t\partial} \}_{t \geq 0} \) with infinitesimal generator \( \partial = \partial_c \) (resp., \( -\partial \)) is a contraction semigroup on \( \mathbb{H}_c \) if \( c \geq 0 \) (resp., \( c \leq 0 \)). Furthermore, if the roles of \( \partial \) and \( -\partial \) (or, equivalently, of \( c \) and \( c + c \)) are interchanged, so are the roles of ‘contraction’ and ‘expansion’ semigroups in this statement. Moreover, for every \( f \in \mathbb{H}_c \) and \( h \in \mathbb{R} \), we have \( \|e^{-t\partial} f\|_c = e^{-|h|} \|f\|_c \) and \( (e^{-t\partial} f)(t) = f(t-h) \), for \( h \in \mathbb{R} \) (so that \( e^{-t\partial} f = f(\cdot - h) \) in \( \mathbb{H}_c \)). As a result, \( \partial \) is also referred to as the infinitesimal shift of the real line.

**Remark 4.3.** Given \( c \in \mathbb{R} \), the c-momentum operator is given by \( p_c := (1/i) \partial_c \), where \( i := \sqrt{-1} \). It is self-adjoint if and only if \( c = 0 \), in which case it coincides with the classic momentum operator \( p_0 \) of quantum mechanics, acting on \( \mathbb{H}_0 = L^2(\mathbb{R}) \) and with spectrum \( \sigma(p_0) = \mathbb{R} \); e.g. [80,107,110].
Similarly, if we let \(-\Delta_c = (p_c)^2 = -(\partial_c)^2\) denote the free \(c\)-Hamiltonian, then the group \(\{e^{-i\theta \Delta_c}\}_{\theta \in \mathbb{R}}\) is unitary if and only if \(c = 0\).

We can next proceed to rigorously define the spectral operator \(a = a_c\) by the expression \(a_c := \zeta(\partial_c)\) or, more concisely, \(a := \zeta(\partial)\), in the sense of the functional calculus for unbounded normal operators (e.g. [104]). Of course, we are using here the fact that (according to theorem 4.1 above), \(\partial_c\) is an unbounded normal operator. Note that, by definition, the domain of \(a\) is given by \(D(a) = \{f \in \mathbb{H}_c : af \in \mathbb{H}_c\}\). In the sequel, \(e^{\ell}(E)\) denotes the closure of \(E \subseteq \mathbb{C}\) in \(\mathbb{C}\).

**Theorem 4.4 ([8]).** For every \(c \in \mathbb{R}\), the spectral operator \(a = a_c\) is a (possibly unbounded) normal operator on \(\mathbb{H}_c\), with spectrum \(\sigma(a)\) given by \(\sigma(a_c) = e^{\ell}(\{\zeta(s) : \Re(s) = c, s \in \mathbb{C}\})\). (When \(c = 1\), the pole of \(\zeta\), we must exclude \(s = 1\) on the right-hand side of the above expression for \(\sigma(a_c)\).) In words, theorem 4.4 states that the spectrum of \(a\) coincides with the closure (i.e. the set of limit points) of the range of \(\zeta\) along the vertical line \(\{\Re(s) = c\}\). This result follows from theorem 4.1 combined with a suitable version of the spectral theorem for unbounded normal operators (see [8, appendix E]) according to which (since \(\zeta\) does not have any eigenvalue and \(\zeta\) is continuous for \(c \neq 1\) (resp., meromorphic for \(c = 1\)) in an open connected neighbourhood of \(\tau(\partial) = \{\tau(s) = c\}\), we have that \(\sigma(a) = \sigma(\zeta(\partial)) = e^{\ell}(\sigma(\zeta(\tau(\partial))))\), for \(c \neq 1\). For \(c = 1\), the extended spectrum is given by \(\sigma(a) := \sigma(a) \cup \{\infty\} = \zeta(\zeta(\partial)),\) where the meromorphic function \(\zeta\) is now viewed as a continuous function from \(\mathbb{C}\) to the Riemann sphere \(\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}\); from which the result also follows for \(c = 1\).

(c) **Justification of the definition of \(\alpha\): quantized Dirichlet series and Euler product**

The first justification for the definition of a given by \(a = \zeta(\partial)\) comes from the following result. In the sequel, \(\mathcal{B}(\mathbb{H}_c)\) stands for the Banach algebra (and, even, \(\mathbb{C}\)-algebra) of bounded linear operators on \(\mathbb{H}_c\), equipped with its natural norm: \(\|L\| := \sup\|Lf\|_c : f \in \mathbb{H}_c, \|f\|_c \leq 1\).

**Theorem 4.5 ([8,12]).** For \(c > 1, a = \zeta(\partial)\) is given by the following operator-valued (or ‘quantized’) Dirichlet series and Euler product, respectively: \(a = \zeta(\partial) = \sum_{n=1}^{\infty} n^{-\partial} = \prod_{p \in \mathbb{P}} (1 - p^{-\partial})^{-1},\) where the series and the infinite product both converge in \(\mathcal{B}(\mathbb{H}_c)\). Furthermore, for any integer \(m \geq 1\) and \(f \in \mathbb{H}_c\), we have \((m^{-\partial})(f)(t) = f(t - \log m),\) for a.e. \(t \in \mathbb{R}\) (and, hence, \((m^{-\partial})(f)\) is \(f(\cdot - \log m)\) in \(\mathbb{H}_c\)).

The latter part of theorem 4.5 follows from proposition 4.2 since \(m^{-\partial} = e^{-(\log m)\partial}\), in the sense of the functional calculus. Naturally, theorem 4.5 also provides a rigorous justification of the heuristic discussion (based on [6,7,§6.3.2]) given in §4a. The following results (also from ([8, ch.7, [12]]) go well beyond that discussion but will not be precisely or fully stated here, by necessity of concision:

1. **(Analytic continuation of \(a_c = \zeta(\partial_c)\), for \(c > 0\)).** For \(c > 0, a = a_c\), coincides with the operator-valued ‘analytic continuation’ of the quantized Dirichlet series and Euler product (defined only for \(c > 1\), initially). More specifically, for a dense subspace of functions \(f\) in \(D(\partial)\) (and hence also in \(\mathbb{H}_c\)), we have, for all \(c > 0, a(f) = (1/(\partial - 1))(f) + \int_0^{\infty} (1/(1 - t^{-\partial}))(f) - t^{-\partial}(f)\) \(dt\), where \([f]\) is the integer part of \(t\). Note that this identity is an operator theoretic version of the analytic continuation of \(\zeta = \zeta(s)\) to \(\{\Re(s) > 0\}\) given in §3a.

2. **(Analytic continuation of \(A_c = \xi(\partial_c)\), for \(c \in \mathbb{R}\), and quantized functional equation.)** Let the global spectral operator be defined by \(A = \xi(\partial)\), where \(\xi\) is the global (or completed) zeta function \(\xi(s) := \pi^{-s/2} \Gamma(s/2) \zeta(s)\). Then, for all \(c \in \mathbb{R}\), the operator-valued analytic continuation of \(A = A_c\) is given by a formula analogous to the one described in (1) above, but in which the symmetry \(s \leftrightarrow 1 - s\) is more immediately manifest. (This is a quantized counterpart of the corresponding formula for \(\xi(s)\); e.g. [33,40] or [20, eqn. (2.4.10)].) One deduces from this identity an operator-valued analogue (for \(A_c\)) of the classic functional equation \(\xi(s) = \xi(1 - s)\). Namely, \(A_c = B_c\), where \(B_c := \xi(I - \partial_c)\). In particular, \(\xi(\partial_{1/2}) = \xi(\partial_{1/2})\).

3. **(Inverse of \(a\) for \(c > 1\)).** For \(c > 1\) (as in theorem 4.5), \(a = \zeta(\partial)\) is invertible (in \(\mathcal{B}(\mathbb{H}_c)\)), with inverse \(a^{-1}\) given by \(a^{-1} = (1/\xi)(\partial) = \sum_{n=1}^{\infty} \mu(n) n^{-\partial},\) where the series is convergent in
We hope that the above discussion provides a sufficiently convincing sample of formulae from quantized number theory, in the terminology of [8–12,28,29]. Many additional formulae can be obtained, involving either ζ(0) (or ξ(0)) or, more generally, ζL(0), where ζL is any of the classic L-functions of number theory and arithmetic geometry; e.g. ([7, appendix A], [20,98,99]). Further exploration in this direction is conducted in [28,29], either in the present framework or else in a different functional analytic and operator theoretic framework (using Bergman spaces, [111,112]); see remark 5.5(b).

(d) Quasi-invertibility of α and the Riemann hypothesis

In 3b above, we have briefly described the work of [4] in which a spectral reformulation of the RH was obtained; see, specifically, theorem 3.4. We will provide here an operator theoretic analogue of this reformulation, based on the work of [8–10].

Given T > 0, let \( a(T) := \xi(\vartheta(T)) \) denote the T-truncated spectral operator, where \( \vartheta(T) := \rho(T)(\vartheta) \) is the T-truncated infinitesimal shift, defined (thanks to theorem 4.1 above) via the functional calculus for unbounded normal operators. In practice, one simply refers to \( a(T) \) and \( \vartheta(T) \) as the truncated spectral operator and the truncated infinitesimal shift, respectively. The precise definition of the cut-off function \( \rho(T) \) is unimportant, provided \( \rho(T) \) is chosen so that the following computation of the spectrum of \( a(T) \) is justified (since \( \sigma(\vartheta) = c + i\mathbb{R} \), by theorem 4.1): \( \sigma(a(T)) = c\ell(\rho(T)(\sigma(\vartheta))) := [c - iT, c + iT] \). Since \( \vartheta(T) = \rho(T)(\vartheta) \), the first equality follows from the spectral mapping theorem (SMT) (applied to the unbounded normal operator \( \vartheta \) and the function \( \rho(T) \); see [8, appendix E]), while the second equality follows from the construction of \( \rho(T) \). In fact, when \( c \neq 1 \), the only requirement about \( \rho(T) \) is that it be a \( \mathbb{C} \)-valued, continuous function on \( c + i\mathbb{R} = [\text{Re}(s) = c] \) whose (closure of the) range is exactly equal to the vertical segment \( [c - iT, c + iT] \). When \( c = 1 \) (which corresponds to the pole of \( \zeta(s) \) at \( s = 1 \)), instead of requiring that \( \rho(T) \) is continuous, we assume that \( \rho(T) \) has a meromorphic continuous to a connected neighbourhood of \( [\text{Re}(s) = c] \). Indeed, when \( c \neq 1 \) (resp., \( c = 1 \)), we can then apply the continuous (resp., meromorphic) version of the SMT given in [8, appendix E].

Again, in light of the functional calculus form of the spectral theorem and of SMT, given in [8, appendix E] (applied to the operator \( \vartheta \) and to the function \( \xi = \xi(s) \), which is continuous along \( \sigma(\vartheta) = [\text{Re}(s) = c] \) if \( c \neq 1 \) and meromorphic in a connected open neighborhood of \( \sigma(\vartheta) \) if \( c = 1 \), the pole of \( \zeta(s) \), we have (as \( a(T) := \xi(\vartheta(T)) \)) that \( \sigma(a(T)) = c\ell(\xi(\sigma(\vartheta(T)))) = c\ell(\xi([c - iT, c + iT])) \). Hence, we conclude (see [8] for the proof) that the spectrum of \( a(T) \) is given by \( \sigma(a(T)) = [\xi(s) : \text{Re}(s) = c, |I_m(s)| \leq T] \), which is a compact subset of \( \mathbb{C} \) for \( c \neq 1 \). For \( c = 1 \), the same formula holds provided one keeps the closure and also requires that \( s \neq 1 \) on the right-hand side of the above expression for \( \sigma(a(T)) \). Alternatively, one can view \( \zeta \) as a continuous \( \mathbb{C} \)-valued function, with \( \bar{\zeta} := \zeta \cup \{\infty\} \), and replace the left-hand side by \( \hat{\sigma}(a(T)) := \sigma(a(T)) \cup \{\infty\} \), the extended spectrum of \( a(T) \). Note that, for \( c \neq 1 \), \( a(T) \) is bounded, so that \( \hat{\sigma}(a(T)) := \sigma(a(T)) \), following the traditional notation.

Remark 4.6. Since, as follows easily from theorem 4.1, we have \( \vartheta = c + iV \), where \( V := (\vartheta - c)/i \) is an unbounded self-adjoint operator with spectrum \( \sigma(V) \) given by \( \sigma(V) = \mathbb{R} \) (and with no eigenvalues), one can equivalently define \( \vartheta(T) \) by \( \vartheta(T) := c + iV(T) \), where \( V(T) := \Phi(T)(V) \) and \( \Phi(T) : \mathbb{R} \to \mathbb{R} \) satisfies \( c\ell(\Phi(T)(\mathbb{R})) = [-T, T] \). In addition, when \( c \neq 1 \), one requires that \( \Phi(T) \) is continuous, while, when \( c = 1 \), one assumes that \( \Phi(T) \) admits a meromorphic extension to a connected neighbourhood of \( \mathbb{R} \) in \( \mathbb{C} \).

We can now introduce the following key definition. The spectral operator \( a \) is said to be quasi-invertible if each of its truncations \( a(T) \) is invertible (for \( T > 0 \)), in the usual sense (and, hence, if
for every $T > 0, 0 \notin \sigma(a(T))$. Using theorem 4.4 and the above expression for $\sigma(a(T))$, along with the fact that an operator is invertible if its spectrum does not contain the origin, it is easy to check that if $a$ is invertible, then it is quasi-invertible. The converse, however, need not be true, as we shall see in §5. (It turns out that assuming that for all $c \in (0, \frac{1}{2})$ the quasi-invertibility of $a = a_c$ is equivalent to its invertibility, is equivalent to the RH; see theorem 5.4.)

The first result in this context is the exact operator theoretic analogue of theorem 3.3 (from §4c).

**Theorem 4.7 (Riemann zeros and quasi-invertibility of $a$, [8,10]).** Fix $c \in \mathbb{R}$. Then, the spectral operator $a = a_c$ is quasi-invertible if and only if $\zeta = \zeta(s)$ does not have any zeros on the vertical line $\{\Re(s) = c\}$.

It follows that, in the midfractal case when $c = \frac{1}{2}, a_{1/2}$ is not quasi-invertible (since $\zeta$ has a zero along the critical line $\{\Re(s) = \frac{1}{2}\}$) and that for $c \geq 0$ (in order to avoid the trivial zeros of $\zeta$, located at $s = -2, -4, \ldots$) $a$ is expected to be quasi-invertible everywhere else. (Recall that, according to Hadamard’s theorem (e.g. [33,40]) and the functional equation for $\zeta$, $\zeta = \zeta(s)$ does not have any zeros along the vertical lines $\{\Re(s) = 1\}$ and $\{\Re(s) = 0\}$.) In particular, we have the following operator theoretic counterpart of theorem 3.4 (from §3b).

**Theorem 4.8 (RH and quasi-invertibility of $a$, [8,10]).** The spectral operator $a = a_c$ is quasi-invertible for all $c \in (0,1)$ (other than for $c = \frac{1}{2}$) (or, equivalently, for all $c \in (0, \frac{1}{2})$ or, else, for all $c \in (\frac{1}{2}, 1)$, respectively) if and only if the RH is true.

Just as for theorem 3.4, we refer to theorem 4.8 as a symmetric criterion for the RH (the symmetry of the roles played by the intervals $(0, \frac{1}{2})$ and $(\frac{1}{2}, 1)$ being due to the functional equation satisfied by $\zeta$). Finally, we close this section by noting that theorem 4.8 also provides a precise and correct functional analytic formulation of the semi-heuristic result stated in Cor. 9.6 of [6,7] about the ‘invertibility’ of $a$.

5. **Invertibility of the spectral operator and an asymmetric reformulation of the Riemann hypothesis**

We have discussed in §3b (based on [3,4]) and in §4d (based on [8–10]) two different, but related (at least in spirit) symmetric reformulations of the RH; see theorem 3.4 (from §3b) and theorem 4.8 (from §4d), respectively. Our goal in this section is to provide a new asymmetric reformulation of the RH (due to the author), as well as to explain why it cannot be modified so as to become symmetric (with respect to the midfractal dimension $c = \frac{1}{2}$), without drastically changing the nature of the problem. As we shall see, this new reformulation is directly expressed in terms of the invertibility of $a$.

In §4d, we have studied the quasi-invertibility of the spectral operator $a = a_c$ (that is, the invertibility of each of its truncations $a^{(T)} = a_{c}^{(T)}$, where $T > 0$ is arbitrary). In fact, both theorems 4.7 and 4.8 (from §4d) are stated in terms of the quasi-invertibility of $a_c$, either for a fixed $c \in (0,1)$, with $c \neq \frac{1}{2}$ (in theorem 4.7), or (in theorem 4.8) for all $c \in (0,1)$, with $c \neq \frac{1}{2}$ (equivalently, for all $c \in (0, \frac{1}{2})$ or, else, for all $c \in (\frac{1}{2}, 1)$). (Recall from the discussion following theorem 4.7 that $a_{1/2}$ is not quasi-invertible and, hence, not invertible.) However, although very convenient in the context of §4d, the new notion of ‘quasi-invertibility’, because it involves a truncation of the infinitesimal shift $\partial$ (and, hence, also of $a$ as well as of its spectrum $\sigma(a)$; see the beginning of §4d), is perhaps not so easy to grasp or to verify explicitly. Indeed, we lack concrete formulae for $a(T)$, whereas we have several explicit expressions for $a$ at our disposal, even when $c \in (0,1)$.

It is therefore natural to wonder whether the (usual) notion of invertibility of the spectral operator $a = a_c$ cannot be used to reformulate the RH in this context. (See the discussion following theorem 4.1 for the usual notion of invertibility of a possibly unbounded operator.) The answer to this question is affirmative, as we shall explain below; see theorem 5.1. Moreover, in light of theorem 4.8 (based on the results of [8–11]), the asymmetry of the resulting criterion for RH turns out to be intimately connected to the universality of the Riemann zeta function $\zeta = \zeta(s)$ in the
right critical strip \( \{ \frac{1}{2} < \text{Re}(s) < 1 \} \) as well as to its (presumed) non-universality in the left critical strip \( \{ 0 < \text{Re}(s) < \frac{1}{2} \} \). This last statement will become clearer as we progress in §5. We can now state and prove the main result of this section. (See also theorem 5.4.)

**Theorem 5.1 (Asymmetric criterion for RH and invertibility of \( a \)).** The spectral operator \( a = a_c \) is invertible for all \( c \in (0, \frac{1}{2}) \) if and only if the RH is true.

By contrast, we have the following result, which was already observed in [8–11] and clearly shows that the reformulation of RH obtained in theorem 5.1 is asymmetric, in a strong sense.

**Theorem 5.2.** The spectral operator \( a = a_c \) is not invertible for any \( c \in (\frac{1}{2}, 1) \).

**Proof of theorems 5.1 and 5.2.**

(i) Let us first prove theorem 5.1. Recall that \( a \) is invertible if and only if \( 0 \notin \sigma(a) \). In light of theorem 4.4, we know that for any \( c \in \mathbb{R} \) (with \( c \neq 1 \)) and with \( a = a_c \), as usual, \( \sigma(a_c) \) coincides with the closure of the range of \( \zeta = \zeta(s) \) along the vertical line (\( \text{Re}(s) = c \)). Now, according to a result of Garunkštis & Steuding in [113] concerning the non-universality of \( \zeta \) in the left critical strip \( \{ 0 < \text{Re}(s) < \frac{1}{2} \} \), we know that, conditionally (i.e. under the RH), we have that, for all \( c \in (0, \frac{1}{2}), \sigma(a) \) (that is, the closure of the range of \( \zeta \) on \( \{ \text{Re}(s) = c \} \) is a strict subset of \( \mathbb{C} \) and, in fact, that \( 0 \notin \sigma(a) \); see [113, lemma 4 and proposition 5] and their proofs. Hence, \( a \) is invertible.

Conversely, assume that \( a = a_c \) is invertible (i.e. \( 0 \notin \sigma(a) \)) for every \( c \in (0, \frac{1}{2}) \). Then, since (by theorem 4.4) \( \sigma(a) \supset \{ \zeta(s) : \text{Re}(s) = c \} \), it follows that \( \zeta(s) \neq 0 \) for all \( s \in \mathbb{C} \) with \( \text{Re}(s) = c \) and every \( c \in (0, \frac{1}{2}) \). In light of the functional equation for \( \zeta \), we then deduce that \( \zeta(s) \neq 0 \) for all \( s \in \mathbb{C} \) with \( 0 < \text{Re}(s) < 1, \text{Re}(s) \neq \frac{1}{2} \); i.e. the RH holds. This concludes the proof of theorem 5.1.

(ii) According to the Bohr–Courant theorem [114], which itself is implied by the universality of \( \zeta \) in the right critical strip \( \{ \frac{1}{2} < \text{Re}(s) < 1 \} \) (see, for example, [115] for an exposition), the range of \( \zeta \) is dense along every vertical line \( \{ \text{Re}(s) = c \} \), with \( \frac{1}{2} < c < 1 \); i.e. in light of theorem 4.4, \( \sigma(a) = \mathbb{C} \) and, hence, \( 0 \notin \sigma(a) \). Therefore, the conclusion of theorem 5.2 holds: \( a \) is not invertible for any \( c \in (0, \frac{1}{2}) \), as desired. This completes the proof of theorem 5.2.

Remark 5.3.

(a) The universality of \( \zeta \) implies a much stronger result than the one used in the proof of theorem 5.2 (part (ii) of the above proof). Namely, for every integer \( n \geq 0 \) and every \( c \in (\frac{1}{2}, 1) \), the range (along the vertical line \( \{ \text{Re}(s) = c \} \) of \( \{ \zeta(s), \zeta'(s), \ldots, \zeta^{(n)}(s) \} \) is dense in \( \mathbb{C}^{n+1} \), where \( \zeta^{(k)}(s) \) denotes the \( k \)th complex derivative of \( \zeta \) at \( s \in \mathbb{C} \) and \( \zeta^{(0)} := \zeta \). This generalization of the Bohr–Courant theorem (from [114]) is due to Voronin in [116] and was probably a key motivation for Voronin’s discovery in [117] of the universality of \( \zeta \) (in the right critical strip). The universality theorem was extended to (suitable) compact subsets of the right critical strip by Bagchi in [118] and Reich in [119]. We refer, for example, to the books [38,115,120] along with [8,11] for many other relevant references and extensions of the universality theorem to other \( L \)-functions.

(b) The universality of \( \zeta \) in the right critical strip roughly means that given any compact subset \( K \) of \( \{ \frac{1}{2} < \text{Re}(s) < 1 \} \) with connected complement in \( \mathbb{C} \), and given any nowhere vanishing (complex-valued) continuous function \( g \) on \( K \) that is holomorphic on the interior of \( K \) (which may be empty), then \( g \) can be uniformly approximated by some (and, in fact, by infinitely many) vertical translates of \( \zeta \) on \( K \).

(c) A quantized (i.e. operator-valued) version of universality is provided in [8,11]. Interestingly, in that context, the natural replacement for the complex variable \( s \) is the family of truncated spectral operators \( \{ \partial(T) \}_{T > 0} \).
The spectral operator $a = a_c$ is bounded (and hence invertible, with a compact spectrum $\sigma(a)$ not containing the origin), for every $c > 1$; it is unbounded (and therefore its spectrum $\sigma(a)$ is a closed unbounded subset of $\mathbb{C}$) for every $c \leq 1$. It is not invertible for $\frac{1}{2} < c < 1$ since $0 \notin \sigma(a) = \mathbb{C}$; by contrast, according to theorem 5.1, it is invertible (i.e. $0 \notin \sigma(a)$) for every $c \in (0, \frac{1}{2})$ if and only if the RH is true. For a discussion of the mathematical phase transition occurring (conditionally) in the midfractal case when $c = \frac{1}{2}$, we refer to [14,15] and [8,9]. In this context of §5, we also refer to the brief discussion at the end of the Introduction.

Let $b = a^* a (= a a^*$, since $a$ is normal). Then $b = b_c$ is a non-negative self-adjoint operator. It is bounded (resp., invertible) if and only if $a$ is; that is, iff $c \geq 1$ (resp., $c > 1$ or, conditionally, $0 < c < \frac{1}{2}$). Furthermore, $b$ is invertible if and only if it is bounded away from zero (i.e. $b \geq \alpha$, for some constant $\alpha > 0$, which may depend on $c$). We may now close this section by stating the following theorem, which provides several reformulations of the RH discussed in this paper, as well as a new one (the last one given in theorem 5.4).

**Theorem 5.4.** The following statements are equivalent:

(i) The Riemann hypothesis is true.

(ii) The spectral operator $a$ is invertible for every $c \in (0, \frac{1}{2})$.

(iii) The spectral operator $a$ is quasi-invertible for every $c \in (0, \frac{1}{2})$ (or, equivalently, for every $c \in (\frac{1}{2}, 1)$).

(iv) For every $c \in (0, \frac{1}{2})$, the unbounded, non-negative, self-adjoint operator $b$ is bounded away from zero (i.e. is invertible).

**Proof.** (i) $\iff$ (ii): This follows from theorem 5.1.

(i) $\iff$ (iii): This follows from theorem 4.8.

(ii) $\iff$ (iv): This follows from the discussion immediately preceding the statement of this theorem.

We note that, as was discussed earlier, the equivalence between (i) and (iii) is a symmetric criterion for RH. By contrast, the equivalence between (i) and (ii) is an asymmetric criterion for RH. Therefore, so is the equivalence between (i) and (iv), which, in practice, might provide the most hopeful way to try to prove the RH along those lines. The author has obtained some preliminary (but not definitive) results in this direction, based on explicit computations and by using a suitable class of ‘test functions’ in $D(b)$, the domain of $b$, but it is premature to discuss them here.

**Remark 5.5.**

(a) It is natural to wonder what is the inverse of $a = a_c$ when $0 < c < \frac{1}{2}$. We conjecture that, under RH, it is given by $a^{-1}(f) = \sum_{n=1}^{\infty} \mu(n)n^{-\alpha} f$ (so that $a^{-1}(f)(t) = \sum_{n=1}^{\infty} \mu(n)f(t - \log n)$), for all $f \in D(a) = \mathbb{H}_c$, where $\mu = \mu(n)$ is the Möbius function (compare with part (3) of §4c). Accordingly, the quantized Riemann zeta function $a = \zeta(b)$ would behave very differently from the ordinary (complex-valued) Riemann zeta function $\zeta = \zeta(s)$. Indeed, even under RH, the Dirichlet series $\sum_{n=1}^{\infty} \mu(n)n^{-s}$ cannot converge for any $s_0 \in \mathbb{C}$ such that $0 < \text{Re}(s_0) < \frac{1}{2}$. If it did, then its sum would have to be holomorphic (see [21] or [22, §VI.2]) and to coincide with $(1/\zeta)(s)$ on $\{\text{Re}(s) > \text{Re}(s_0)\}$, which is impossible because $1/\zeta$ must have a pole at every zero of $\zeta$ along the critical line $\{\text{Re}(s) = \frac{1}{2}\}$. (See also, for example, [33].)

(b) The same methods as those discussed in this paper can be applied to a very large class of $L$-functions (or arithmetic zeta functions), all of which are expected to satisfy the GRH (e.g. [7, appendix A], [20, appendices B, C and E], [98,99]). For obtaining the analogue of the results of §5, we would need to use, in particular, the universality results discussed in [115]. The relevant results of [113] used in the proof of theorem 5.1 would also need to be extended to this broader setting.

(c) In [28,29], the author has adapted (and extended in various directions), the present theory to a new functional analytic framework (based on weighted Bergman spaces of...
At this early stage, it seems that, depending on the goal being pursued, one approach or the other may be more appropriate. In particular, the approach discussed in this article (beginning with §4) is ideally suited for formulating and obtaining the new results of §5.

Competing interests. I declare I have no competing interests.

Funding. This research was partially supported by the US National Science Foundation (NSF) under grants DMS-0707524 and DMS-1107750 (as well as by many earlier NSF grants since the mid-1980s).

References


**Glossary**

- *abscissa of convergence*, 4
- \( a_k \sim b_k \) as \( k \to \infty \), asymptotically equivalent sequences, 8
- \( \overline{c(A)} \), closure of a subset \( A \) in \( \mathbb{C} \) (or in \( \mathbb{R}^N \)), 13
- \( \alpha = \alpha_c \), spectral operator, 11, 13
- \( \alpha^{(T)} = \alpha^{(T)}_c \), truncated spectral operator, 14
b = aa^* = a^a, where a is the spectral operator, 17

complex dimensions (of a fractal string), 5

CS, Cantor string, 6

\( \mathbb{C} \), the field of complex numbers

\( \hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\} \), the Riemann sphere, 13

direct spectral problem, 8–9

Dirichlet series, 5, 13

D, Minkowski dimension of a fractal string (or of \( A \subset \mathbb{R}^N \)), 4

\( D = D_L \), the set of complex dimensions of a fractal string \( L \) (or of its boundary), 5

\( \text{dist}(x, A) = \inf\{|x - a| : a \in A\} \), Euclidean distance from x to \( A \subset \mathbb{R}^N \), 6

\( \partial = \partial_c \), infinitesimal shift of the real line, 12

\( \partial(T) = \partial_c(T) \), truncated infinitesimal shift, 14

\( \partial \Omega \), boundary of a subset \( \Omega \) of \( \mathbb{R}^N \) (or of a fractal string), 4

Euler product, 5, 13

\( |E| = |E|_N \), N-dimensional Lebesgue measure of \( E \subset \mathbb{R}^N \), 4 (\( N = 1 \)), 9

general explicit formulae, 7

\( I'(t) := \int_0^{+\infty} x^{t-1} e^{-x} \, dx \), the gamma function, 5

\( \mathbb{H}_c := L^2(\mathbb{R}, e^{-2ct} \, dt) \), weighted Hilbert space, 12

\( i = \sqrt{-1} \), imaginary unit, 6

(ISP)_D, inverse spectral problem for the fractal strings of Minkowski dimension D, 9–10

invertibility, 12

\( \xi = \xi(s) \), completed (or global) Riemann zeta function, 5

\( \log_a x \), the logarithm of \( x > 0 \) with base \( a > 0 \); \( y = \log_a x \Leftrightarrow x = a^y \), 6

\( \log x \), the natural logarithm of \( x \); \( y = \log x \Leftrightarrow x = e^y \), 6

\( \mathcal{L} = \{\ell_j\}_{j=1}^{\infty} \), a fractal string with lengths \( \ell_j \), 5

midfractal case, 1, 4, 10, 15, 17

Minkowski content (of a Minkowski measurable fractal string), 4

Minkowski non-degenerate (fractal string), 4, 9

Minkowski measurable (fractal string), 4, 8

\( \mathcal{M}_{s,d} \) and \( \mathcal{M}^*_d \), lower and upper \( d \)-dimensional contents of a fractal string \( L \) (or of a bounded set \( A \subset \mathbb{R}^N \)), 4, 9

\( \mathcal{M}, \mathcal{M}_s \) and \( \mathcal{M}^* \), Minkowski content, lower and upper Minkowski content (of \( L \) or of \( A \subset \mathbb{R}^N \)), 4, 9

\( \mu(n) \), Möbius function, 13–14, 17

\( N_L \), geometric counting function of a fractal string \( L \), 5, 8–9

\( N_\nu \), spectral (or frequency) counting function of a fractal string or drum, 5

\( \Omega_\varepsilon \), \( \varepsilon \)-neighbourhood of a fractal string (or of a fractal drum), 4

\( \mathcal{P} \), the set of prime numbers, 5

quasi-invertibility, 14

\( \text{res}(f; \omega) \), residue of the meromorphic function \( f \) at \( \omega \in \mathbb{C} \), 6

Riemann’s explicit formula, 6

Riemann hypothesis, 2, 10, 15, 17

\( \sigma(L) \), spectrum of a fractal string \( L \), 5

\( \sigma(T) \), spectrum of the operator \( T \), 12–13

\( \tilde{\sigma}(T) \), extended spectrum of the operator \( T \), 13
V(ε), volume of the ε-neighbourhood of a fractal string (or drum), 4

W(x), Weyl term, 8–9

[x], integer part of x ∈ ℝ, 8

ζ(s) = ∑ j=1 j−s, Riemann zeta function, 4–5

ζ_L(s) = ∑ j=1 (ℓ_j)^s, geometric zeta function of a fractal string L = {ℓ_j}^∞, 4–5

ζ_v = ζ_{v,L}, spectral zeta function of a fractal string L, 4–5

\|·\|_c, norm of the Hilbert space H_c, 12

\langle·, ·\rangle_c, inner product of the Hilbert space H_c, 12