Overview

In Chapter 1, we give a broad introduction to several of the main themes encountered in this work: arithmetic geometry, noncommutative geometry, quantum physics and string theory, prime number theory and the Riemann zeta function, along with fractal and spectral geometry.

In Chapter 2, we explain how string theory on a circle (or on a finite-dimensional torus)—considered from the point of view of Connes' noncommutative geometry, as in the work of Fröhlich and Gawędzki, pursued by Lizzi and Szabo—can be used as the starting point for a geometric and physical model of the Riemann zeta function ζ and other arithmetic L-series. In particular, by analogy with the key role played by the Poisson Summation Formula in both the physical and the arithmetic theory, we contend that the classic functional equation satisfied by ζ corresponds to T-duality in string theory. The latter, a key symmetry that is not present in ordinary quantum mechanics, allows one to identify physically and mathematically two circular spacetimes with reciprocal radii. Furthermore, we suggest that the Riemann Hypothesis may be related to the existence of a fundamental length in string theory.

In Chapter 3, we first briefly review some aspects of the author's theory of fractal strings (one-dimensional drums with fractal boundary) and of the associated theory of complex dimensions, as developed in the research monograph [Lap-vF2] (joint with M. van Frankenhuysen) *Fractal Geometry and Number Theory*: *Complex dimensions of fractal strings and zeros of zeta functions* (Birkhäuser, Boston, 2000). [See also the new book [Lap-vF9], *Fractal Geometry, Complex Dimensions and Zeta Functions*: *Geometry and spectra of fractal strings* (Springer-Verlag, New York, 2006).] We then introduce the new concept of a fractal membrane, a suitable multiplicative (or quantum) analogue of a fractal string. Heuristically, a fractal membrane can be thought of as a (noncommutative) Riemann surface with infinite genus or as an (adelic) infinite dimensional torus. We show that the (spectral) partition function of a fractal membrane is naturally given by an Euler product, which reduces to the usual one for ζ in the case of the 'prime membrane' associated with the Riemann zeta function (or, equivalently, with the field of rational numbers). We thus obtain in this case a new mathematical model (different from that of Bost and Connes) for the notion of a 'Riemann gas' introduced by the physicist B. Julia in the context of quantum statistical physics. We point out, however, that our motivations and goals in developing the theory of fractal membranes are significantly broader than in the latter work, as is discussed in parts of Chapters 4 and 5.

Towards the end of Chapter 3, we also introduce the new (but closely related) concept of self-similar membrane, which corresponds to a different choice of statistics than for a fractal membrane when quantizing a fractal string. In a special case, the spectral partition function of a fractal membrane is shown to coincide with the geometric zeta function of a self-similar fractal string. By comparing our notions of fractal and self-similar membranes, we also develop a useful parallel between aspects of arithmetic and self-similar geometries. We strengthen this analogy and close Chapter 3 by providing a dynamical interpretation of the partition functions of fractal membranes and of self-similar membranes. In the former case, the associated suspended flows may be called 'Riemann–Beurling flows'. Indeed, the logarithms of the underlying (generalized) primes coincide with the 'weights' (or 'lengths') of the corresponding primitive orbits. We note that in our context, the 'Riemann flow' is associated with the 'prime fractal membrane' (or, equivalently, with the field of rational numbers).

In Chapter 4, we discuss various noncommutative and increasingly rich models of fractal membranes. In particular, we briefly discuss some very recent work of the author (joint with R. Nest) in which we show that fractal (and self-similar) membranes are the second (or Dirac) quantization of fractal strings. In this context, the choice of Fermi–Dirac—or Bose–Einstein, in a second and improved construction—quantum (resp., Gibbs) statistics corresponds to fractal (resp., self-similar) membranes. In short, it follows that fractal membranes (or their self-similar counterparts) can truly be considered as 'quantum fractal strings'. One of the new heuristic and mathematical insights provided by the latter work is that once fractal strings have been quantized, their endpoints are no longer fixed on the real axis but are allowed to move freely within suitable copies of the holomorphic disc in the complex plane. This seems to be somewhat analogous to the role played by D-branes in contemporary string theory or in M-theory.

As is explained earlier on in Chapter 3, one can associate a prime fractal membrane to each type of arithmetic geometry, including algebraic number fields and function fields (for example, curves or higher-dimensional varieties over finite fields). Near the end of Chapter 4, we propose that a more geometric, algebraic and physical model of arithmetic geometries can be based on the 'noncommutative stringy spacetime' corresponding to closed strings propagating in a fractal membrane—viewed, for example, as an adelic infinite dimensional torus. Such a spacetime can be thought of as a sheaf of ordinary noncommutative or quantum spaces—and thus, in our framework, of vertex operator algebras along with dual (or 'chiral') pairs of Dirac operators. The functional equation satisfied by an arithmetic zeta function such as ζ would then be the analytic counterpart of Poincaré duality at the cohomological level, and of T-duality, at the physical level. Accordingly, we conjecture that a suitable spectral and cohomological interpretation of the (dynamical) complex dimensions of fractal membranes—and, in particular, of the 'Riemann zeros', i.e., the nontrivial zeros of ζ—can be obtained in this context, by means of the associated sheaf of vertex algebras.

In Chapter 5, we suggest that the author's moduli spaces of fractal strings and of fractal membranes—viewed as highly noncommutative spaces significantly generalizing the set of all Penrose tilings—should be a natural receptacle for zeta functions and for a suitable extension of Deninger's heuristic notion of 'arithmetic site'. We conclude by proposing a new geometric and dynamical interpretation of the Riemann Hypothesis, expressed in terms of a suitable noncommutative flow of zeta functions acting on the moduli space of fractal membranes, along with the associated flow of zeros. (Each of these flows is referred to as a 'modular flow' or as an 'extended Frobenius flow'.)

Accordingly, conjecturally, along the orbits of the modular flow of fractal membranes, the associated generalized, noncommutative fractal geometries would be 'continuously deformed' (i.e., would 'converge') to arithmetic geometries—viewed as stable, attractive 'fixed points' of this noncommutative flow. Consequently, the truth of the Riemann Hypothesis (and of its natural extensions) would follow from the convergence of the zeros of the corresponding zeta functions to the critical line—or, equivalently, to the Equator of the Riemann sphere, both from within the lower and upper hemispheres, using T-duality and the associated 'generalized functional equations'.

We close Chapter 5 by drawing analogies between our conjectural 'modular flows of zeta functions and of their associated zeros' and other flows arising naturally in contemporary mathematics and physics. These flows include Wilson's renormalization flow, the Ricci flow on three-dimensional manifolds, as well as the 'KP-flow' (viewed as a noncommutative, geodesic flow). Accordingly, our modular flow of zeta functions could perhaps be viewed as a noncommutative and arithmetic analogue of the Ricci flow. Similarly, the associated flow of zeros could be thought of as an arithmetic, noncommutative KP-flow. In this chapter, we also propose a model of our modular flows, which is called the 'KMS-flow' (for generalized Pólya-Hilbert operators) and is motivated in part by analogies with quantum statistical physics (in the operator algebraic formalism), along with the Feynman integral and renormalization flow (or group) approaches to quantum systems with highly singular interactions.

It may be useful for the reader to be aware from the outset of the following distinction between the various parts of this book. While Chapter 1 is intended for a 'general' scientific audience, Chapter 2 is more physics-oriented (but still accessible to mathematicians not familiar with string theory), whereas the rest of the book (Chapters 3–5) is clearly of a much more mathematical nature, even though in various places it draws on the physical language, intuition and formalism discussed in Chapter 2. Relevant background material is provided in several places within the text, as well as in the six appendices, in order to make the book more easily accessible and facilitate the transition between its various parts.

As was just mentioned, we have tried to write this book in such a way that someone not familiar with all the subjects dealt with here can still understand the main ideas and concepts involved. We should caution the reader, however, that the mathematics underlying parts of the theory presented in this work is rather formidable and, in fact, is often not yet fully developed or even precisely formulated. We hope, nevertheless, that our proposed models may provide a useful bridge between various aspects of noncommutative, string, arithmetic and fractal geometry as well as, in the long term, motivate further investigations aimed at understanding the elusive geometry underlying the prime numbers (or the integers) and the Riemann zeros.

CHAPTER 1

Introduction

The recipes for quantization are a primitive manifestation of the fact that the space of internal degrees of freedom "at a single point" *in vacuo* is already infinite dimensional because of the virtual generation of particles. Further understanding is blocked until we relinquish the idea of space-time as the basis for all of physics.

Yuri I. Manin, 1979 [Mani1,p.94]

One would of course like to have a rigorous proof of this, but I have put aside the search for such a proof after some fleeting vain attempts because it is not necessary for the immediate objective of my investigation.

Bernhard Riemann, 1859 [Rie1], introducing his famous "Riemann Hypothesis". (Translated in [Edw,p.301].)

1.1. Arithmetic and Spacetime Geometry

I believe that at its deepest level, the geometry underlying the integers—in the old language, the 'geometry of numbers', and in modern terminology, 'arithmetic geometry', including the twin mystical notions of the 'arithmetic site' [Den3,6;Har2] and of the 'field of one element' [Mani4;So1,3]—would have to reflect the physical and geometrical properties of what we traditionally call 'spacetime', for lack of a better word.

I have held this belief, at least consciously, since the mid-1980's when I read the beautiful paper by Yuri I. Manin [Mani2], entitled *New Dimensions in Geometry*[1]. It was later strengthened and turned into an intimate conviction by my own reflections and research experiences in developing the theory of 'fractal strings'[2] since the late 1980's and exploring its relationships with aspects of number theory, particularly the Riemann zeta function and the Riemann Hypothesis [Lap1–4,LapPo1–3,LapMa1–2, HeLap1–2,Lap-vF1–5,9]. In July 1994, I was startled to hear Alain Connes express a similar belief during a debate held at UNESCO in Paris on the occasion of the International Congress of Mathematical Physicists. From our ensuing conversations about this subject—and from our ongoing dialogue (since the summer of 1993) about our respective approaches to the Riemann zeta

[1] I am grateful to Christophe Soulé for sharing with me his enthusiasm for this paper and for Arakelov theory [SoABK] when I first met him in Berkeley in August 1984.

[2] or 'fractal harps', as sometimes referred to in [Lap-vF2,9], not to be mistaken with the strings encountered in the classical string theory [Del3,GreSWit,Gree,Kak,Mani3,Polc3–4,Schw1], although part of the point of the present book is that the two theories can be related, albeit in unexpected ways.

function ([BosCon1–2], surveyed in [Con6,§V.11], and [LapPo1–2,LapMa1–2,Lap2–4,HeLap1–2], now pursued in [Lap-vF1–5,9])—it appears, however, that his vision then (and probably still now in his new approach [Con9,10]) is quite different from the model I am about to propose, although key aspects of noncommutative geometry play an important role in both cases.[3]

1.2. Riemannian, Quantum and Noncommutative Geometry

During the course of the 20th century, and ever since the resounding success of the application of Riemannian geometry to the study of gravity in Einstein's theory of general relativity, geometry has been a focal point for many mathematicians and physicists interested in apprehending aspects of physical reality. As is well known, symplectic geometry is well suited to and, in fact, largely motivated by the study of phase space in classical mechanics. Furthermore, as was mentioned just above, Riemannian geometry—in its Lorentzian version—is adopted in most models of classical physics concerned with gravitational fields. More recently, the geometry (and topology) of principal bundles over differentiable manifolds has been found to be an ideal tool to explore gauge field theories. Note, however, that most mathematically rigorous investigations of gauge theory to date have focused on classical rather than quantum aspects.

It is much less clear, at the moment, how to determine what is "the" geometry underlying quantum mechanics, let alone quantum field theory. More generally, we do not understand what are the true mathematical foundations of quantum field theory [Wit17,19]. Of course, this question has been the subject of much speculation and controversy. In recent years, noncommutative geometry has arisen in large part as a possible answer to such a question, although it is fair to say that we still seem to be far from having resolved this crucial problem. Beginning with the algebraic and functional analytic work of Murray and von Neumann [Mu-vN,vN], as well as of Gel'fand and Naimark [GelfNai], noncommutative geometry truly emerged and flourished as an independent subject with the deep work of Alain Connes. (See, for example, the books [Con5] and [Con6]; see also [GraVarFi].) In essence, the central objects of noncommutative geometry are no longer spaces of points, as in ordinary geometry, but (typically noncommutative) operator algebras, the elements of which can be thought of heuristically as representing quantum fields on the underlying 'noncommutative (or quantum) spaces'. In recent years, Connes [Con7,8] has proposed a set of axioms for noncommutative geometry that requires a much richer structure for a noncommutative space. It involves, in particular, the existence of a suitable Dirac-type operator acting on the Hilbert space on which the operator algebra is represented. (Intuitively, the noncommutative algebra itself can be thought of as the 'algebra of coordinate functions' on the associated quantum space.) This enables one, for example, to measure distances within a noncommutative space much as in a Riemannian manifold, using a formula in some sense dual to the geodesic formula. (See [Con4] and [Con6,Chapter VI].) Under appropriate assumptions, this also provides a noncommutative analogue of the de Rham complex and of aspects of differential topology and geometry (see [Con2–8] and [GraVarFi]). It is good to keep in mind, as is often stressed by Daniel Kastler [Kast2,3], that the aforementioned axioms are largely motivated by models from quantum physics, particularly the so-called 'Standard Model' for elementary particles (see [DV-K-M],

[3]See, however, the relevant discussion in §5.4 for some possible connections.

[ConLo], [Con5,6]), as well as by the long-standing problem of quantum gravity (see, e.g., [ChaCon1,2]).

1.3. String Theory and Spacetime Geometry

Over the last twenty years, string theory, which originated as a theory of strong interactions in the early 1970's, soon to be superseded by quantum chromodynamics (QCD), has emerged as the best candidate for unifying the four known fundamental forces (or interactions) of nature: the electromagnetic force, the weak force and the strong force—all described by Yang–Mills gauge field theories—along with the gravitational force, described by Einstein's theory of general relativity. In this sense, it may eventually provide a means of fully reconciling quantum mechanics (or quantum field theory) with general relativity, and thereby resolve the riddle posed by quantum gravity. Caution must be exercised, however, because despite its great beauty and mathematical power, string theory is still far from being a complete physical or mathematical theory. Moreover, due to the extremely high energies (or, equivalently, the minuscule scales) involved, it has been notoriously difficult in string theory to make predictions that can be verified experimentally with the technology available at present or even in the foreseeable future. We note, however, that although experiments involving high-energy accelerators seem to be out of the question—except to verify some of the most basic assumptions of (super)string theory, such as the existence of supersymmetry [Kan,Freu,Wein4]—interesting large-scale astronomical experiments currently under way may provide useful clues within the next ten to fifteen years. It is also worth mentioning that very recently, low-energy experiments in nuclear and condensed matter physics have confirmed the existence of the so-called 'dynamical supersymmetry' for heavy nuclei (see [Is], [Jol]), but cannot be regarded as providing conclusive evidence for supersymmetry in fundamental physics, while experimental tests for the existence of extra dimensions of spacetime (as required by string theory) have been proposed for the next generation of high-energy accelerators (see, e.g., [Ant]).

Roughly speaking, in string theory, point-particles are replaced with tiny strings (i.e., one-dimensional open strings or else closed loops) vibrating in a (target) spacetime, which is assumed to be ten-dimensional in superstring theory. As it evolves with time, a given string sweeps out a two-dimensional world-sheet, viewed mathematically as a Riemann surface. Hence, the Feynman path integral approach to quantum mechanics ([Fey1], [FeyHi], see also [JohLap]) naturally extends to this setting, with the path integral being replaced by an integral over all possible world-sheets, or more precisely, with integrals over suitable moduli spaces of Riemann surfaces (with a given finite genus and a given number of marked points). The resulting heuristic Feynman-type integral is often referred to as a 'Polyakov integral' [Poly1–3] in the literature. (See, for example, [GreG,GreSWit,Kak,Polc3–4,Wit4], along with [JohLap], Chapter 20, especially Section 20.2.B.) The associated Feynman (or string) diagrams take a much simpler form than in quantum field theory and their detailed analysis provides a good understanding of perturbative string theory, at least at the physical level of rigor. The miracle is that the divergences caused by the coincidence of points in spacetime (and hence the vanishing size of point-particles) in standard quantum field theory now disappear because of the extended size of the strings. In physical terms, superstring theory is said to be renormalizable or, more precisely, "finite to all orders in perturbation theory."

In concluding his plenary lecture at the International Congress of Mathematicians delivered in Berkeley in 1986, Edward Witten made the following statement ([Wit4,p.302],1987):

> *I have tried to make it plausible that path integrals on Riemann surfaces can be used to formulate a generalization of general relativity. What is more, the resulting generalization is* (*especially in its supersymmetric forms*) *free of the ailments that plague quantum general relativity. If the logic has seemed a bit thin, it is at least in part because almost all we know in string theory is a trial and error construction of a perturbative expansion.* [The Feynman–Polyakov path integrals over moduli spaces of Riemann surfaces] *are probably the most beautiful formulas that we now know of in string theory, yet these formulas are merely a perturbative expansion … of some underlying structure. Uncovering that structure is a vital problem if ever there was one.*

Such was the situation up to the late 1980's. However, during the 1990's, significant progress was made towards developing a nonperturbative string theory, called M-theory, in which (one-dimensional) strings are replaced with higher-dimensional geometric objects, called 'membranes' or 'D-branes'. The associated 'dualities' (including the so-called 'S-duality' and 'T-duality') enable one to relate the five basic types of string theory,[4] and thereby to obtain a more unified picture of string theory. (See, for example, [Wit15–17] and [GivePR,Gree,Polc1–4,Schw2–4,Va1–2].) These recent developments are sometimes referred to as the "second superstring revolution" [Schw2].

Edward Witten often begins his lectures on string theory—especially when addressing a mathematical audience—by stressing a striking contrast between the historical developments of string theory and general relativity (see also, for example, the introduction of [Wit4]). In ([Wit13,pp.205–206],1994), he writes:

> *More fundamentally, I believe that the main obstacle* [to further progress] *is that the core geometrical ideas—which must underlie string theory the way Riemannian geometry underlies general relativity—have not yet been unearthed.*

Whatever the true underlying geometric foundations of string theory (or of M-theory), there seems to be an emerging consensus among theoretical and mathematical physicists that one needs to significantly revise the notion of spacetime, from both geometrical and physical points of view. In particular, at extremely small scales (typically, below the Planck scale[5]), the classical model of spacetime as a smooth Riemannian (or Lorentzian) manifold is probably no longer valid. For

---

[4] One of these, the so-called (standard) superstring theory, lives in a ten-dimensional spacetime, consisting of three plus one extended space and time dimensions along with six 'compactified' (tiny) space dimensions.

[5] The *Planck length* (or scale) is the fundamental scale of quantum gravity. It is approximately equal to 1.6×10^{-33} cm (in international units) and is expressed in terms of the following three universal constants, $\hbar$ (Planck's constant or quantum of action), c (the speed of light), and G (Newton's gravitational constant). It is also equal to the reciprocal of the *Planck mass*, about 1.22×10^{19} GeV, the natural mass (or energy) scale of quantum gravity. (It may be useful to note—as is frequently stressed by physicists—that the Planck length is about 20 orders of magnitude

example, the small-scale structure of spacetime may be discrete, or partly discrete and partly smooth. Alternately, it may be of a fractal nature. In fact, in early work on quantum gravity by Wheeler [Whe,WheFo], Hawking and others (see, for example, [GibHaw,Haw,HawIs]), there were intriguing references to the existence of some kind of 'fractal foam', sometimes also called 'quantum foam'. (More recently, see also [Not] in another context.) More radically, it has even been suggested that we do away with the notion of spacetime altogether, at least as a primary concept. (See, for instance, Witten's article [Wit15] entitled *Reflections on the Fate of Spacetime*, from which the second quote heading this book is excerpted. Also, for a different perspective on a similar theme, see Manin's quote from [Mani1] heading the present introduction.) Perhaps an appropriate modification or extension of Connes' noncommutative geometry [Con5,6] will provide clues as to how to proceed in suitably altering or replacing the concept of spacetime. Indeed, there has already been a number of attempts in this direction, several of which will be key to aspects of our present work. (See, for example, [Wit3], and more recently, [FroGa,ChaFro,Cha1–2,LiSz1–2,FroGrRe1–2] along with [ConDouSc,LanLiSz].)

Whatever the answers to these fundamental questions ultimately turn out to be, the relationship between physics and geometry (in a broad sense) will continue to be at the center of the ongoing dialogue between physicists and mathematicians during the next few decades of the 21st century.

It may be helpful at this stage to briefly explain in physical terms the role played by the vibrations of strings in superstring theory (the marriage of string theory and supersymmetry). In quantum field theory (QFT), elementary particles—or rather particle types, such as photons, electrons, quarks, etc.—are represented as *quantum fields* (mathematically, suitable operator-valued distributions; physically, "*bundles*" *or quanta of* "*energy and momentum*" [Wein5,pp.96–97]).[6] In string theory, however, they appear as *the different modes of vibration of the* (closed or open) *strings* "*that make-up the fabric of spacetime*" [Wein5]. (See also [Gree].) At sufficiently low energy, superstring theory can be shown to yield quantum field theory (which is therefore referred to as an *effective theory*). More specifically, the Standard Model of elementary particles [Wein1–5] can be recovered as a low-energy approximation of superstring theory [Del3,Polc3-4]. For example, one of the modes of string vibration corresponds to a particle of spin 1 and zero mass, namely a photon, the carrier (or quantum) of electromagnetic interactions in quantum electrodynamics (QED). Moreover, another mode of string vibration corresponds to a particle of spin 2 and zero mass, which is identified with the graviton, the (presumed) quantum of the gravitational field. In this sense, superstring theory enables us to quantize general relativity (Einstein's theory of gravitational interactions). In fact, as is stressed by Steven Weinberg in his stimulating essays [Wein1,5], "*string theories not only unite gravitation with the rest of elementary particle physics, they explain why gravitation must exist*" [Wein5,p.65].

In order for quantum gravity or the Standard Model to be well understood in the context of string theory, we will still have to overcome formidable obstacles.

smaller than the size of a proton.) In much of this work, we will choose units so that the Planck length (or rather, the string length, see §2.2.3) is equal to one.

[6] More accurately, in quantum field theory, quantum fields are the '*primary concepts*', whereas particles are only '*derived concepts*'—see [Wein5] along with, e.g., [Wein2–4].

For example, long-standing open questions such as understanding the specific numerical values and the wide range of the masses of the elementary particles and of the strengths (or 'coupling constants') of the fundamental interactions seem to be completely out of reach for the time being (see, e.g., [Wit19]), and may remain without any satisfactory answer for a long time to come. Fortunately, these are problems beyond the scope of the present book.

1.4. The Riemann Hypothesis and the Geometry of the Primes

> The theory of Numbers has always been regarded as one of the most obviously useless branches of Pure Mathematics. The accusation is one against which there is no valid defence; and it is never more just than when directed against the parts of the theory which are more particularly concerned with primes. A science is said to be useful if its development tends to accentuate the existing inequalities in the distribution of wealth, or more directly promotes the destruction of human life. The theory of prime numbers satisfies no such criteria. Those who pursue it will, if they are wise, make no attempt to justify their interest in a subject so trivial and so remote, and will console themselves with the thought that the greatest mathematicians of all ages have found in it a mysterious attraction impossible to resist.
>
> ... Very different results are revealed when we turn to the second principal branch of the modern theory, the theory of the *average or asymptotic distribution of primes.* This theory (though one of its most famous problems is still unsolved) is in some ways almost complete, and certainly represents one of the most remarkable triumphs of modern analysis. The theory centres around one theorem, the *Primzahlsatz* or *Prime Number Theorem*; and it is to the history of this theorem, which may almost be said to embody the history of the whole subject, that I shall devote the remainder of this lecture.
>
> ... The next great step was taken by Riemann in 1859, and it is in Riemann's famous memoir *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse* that we first find the ideas upon which the theory has now been shown really to rest. Riemann did not prove the Prime Number Theorem: it is remarkable, indeed, that he never mentions it. His object was a different one, that of finding an explicit expression for $\pi(x)$ [the number of primes not exceeding x, denoted by $\Pi(x)$ in this book], or rather for another closely associated function, as a sum of an infinite series. But it was Riemann who first recognized that, if we are to solve any of these problems, we must study the Zeta-function as a function of the *complex* variable $s = \sigma + it$, and in particular study the distribution of its zeros.
>
> ... To these propositions [Riemann] added certain others of which he could produce no satisfactory proof. In particular he asserted that there is in fact an infinity of complex zeros, all naturally situated in the 'critical strip' $0 \leq \sigma \leq 1$; an assertion now known

to be correct. Finally he asserted that it was 'sehr wahrscheinlich' [very probable] that all these zeros have the real part $\frac{1}{2}$: the notorious 'Riemann hypothesis', unsettled to this day.

We come now to the time when, a hundred years after the conjectures of Gauss and Legendre [about the asymptotic distribution of the primes], the theorem was finally proved. The way was opened by the work of Hadamard on integral transcendental functions. In 1893 Hadamard proved that the complex zeros of Riemann actually exist; and in 1896 he and de la Vallée–Poussin proved independently that *none of them have the real part* 1, and deduced a proof of the Prime Number Theorem.

It is not possible for me now to give an adequate account of the intricate and difficult reasoning by which these theorems are established. But the general ideas which underlie the proofs are, I think, such as should be intelligible to any mathematician.

. . . The arguments which I have advanced are not exact: I have merely put forward a chain of reasoning which seems likely to lead to the desired result. The achievement of Hadamard and de la Vallée–Poussin was to replace these plausibilities by rigorous proofs. It might be difficult for me to make clear to you how great this achievement was. Some branches of pure mathematics have the pleasant characteristic that what seems plausible at first sight is generally true. In this theory anyone can make plausible conjectures, and they are almost always false. Nothing short of absolute rigour counts; and it is for this reason that the Analytic Theory of Numbers, while hardly a subject for an amateur, provides the finest possible discipline in accurate reasoning for anyone who will make a real effort to understand its results.

Godfrey H. Hardy, 1915 [Hard2,pp.350–354],
in his lecture on *Prime Numbers*

The zeta-function is probably the most challenging and mysterious object of modern mathematics, in spite of its utter simplicity.

. . . The main interest comes from trying to improve the Prime Number Theorem, i.e., getting better estimates for the distribution of the prime numbers. The secret to the success is assumed to lie in proving a conjecture which Riemann stated in 1859 without much fanfare, and whose proof has since then become the single most desirable achievement for a mathematician.

Martin C. Gutwiller, 1990 [Gut2,p.308]

The Riemann Hypothesis would say that looking for primes is rather like tossing a coin. [. . .] Riemann predicted that the error term in [the Prime Number Theorem] is the same as the error we expect to see when tossing coins, making primes look in some sense like a random process. [This] distribution of the primes conjectured by Riemann is as nice as we could hope for.

M. *du Sautoy*, 1998 [dSa]

It is perhaps fitting that the same mathematician who brought Riemannian geometry to the world (with such an impact on physics, especially general relativity, half a century later) also proposed what later came to be known as the most famous open problem of mathematics, the so-called Riemann Hypothesis. In developing his geometry, Georg Friedrich Bernhard Riemann (1826–1866) was motivated by the work of his predecessors—including Karl Friedrich Gauss (1777–1855) and the co-discoverers of non-Euclidean geometry, Nikolai Ivanovich Lobachevsky (1792–1856) and Johann (or János) Bolyai (1802–1860)—as well as by philosophical and physical considerations.[7] On the other hand, Riemann's Conjecture (or Hypothesis) concerning the location of the critical zeros of the Riemann zeta function $\zeta = \zeta(s)$—namely, $\zeta(s) = 0$ with $0 \leq \text{Re } s \leq 1$ implies that $\text{Re } s = \frac{1}{2}$—seems to have had entirely different and purely 'internal' (hence, mathematical) motivations.

The Riemann Hypothesis has fascinated mathematicians since its introduction by Riemann in his famous inaugural lecture to the Berlin Academy of Sciences in 1859 (see [Rie1]). Curiously, it was presented almost as a passing remark or conjecture within [Rie1], the only paper by Riemann devoted to number theory. (See the second quote heading this introduction.) Although never stated overtly, one of the main goals of Riemann in [Rie1] seems to have been to provide the tools needed to establish the (then still unproven) 'Prime Number Theorem' conjectured by Gauss and Legendre, according to which, in particular,

$$\Pi(x) = \frac{x}{\log x}(1 + o(1)) \tag{1.4.1}$$

as $x \to \infty$, where the symbol $o(1)$ denotes a function tending to zero as $x \to \infty$ and $\Pi(x) = \Sigma_{p \leq x} 1$ denotes the 'prime number counting function', equal to the number of primes p not exceeding $x > 0$. The Prime Number Theorem[8] was eventually proved almost forty years later in 1896, simultaneously and independently by Jacques Hadamard [Had2] and Charles-Jean de la Vallée Poussin [dV1]. (See also the earlier key papers [vM1,2] and [Had1], along with the later and more precise error estimate obtained in [dV2].) We refer the interested reader to Edwards' book [Edw] or to W. Schwarz's recent survey article [Schwa] for a detailed history of the Prime Number Theorem. As is well known (see, for example, [Edw], [In], or [Pat,§1.8]), the Riemann Hypothesis is equivalent to the statement that the prime numbers are asymptotically distributed as 'harmoniously' as possible or, more precisely, that the error term in the statement of the Prime Number Theorem (in the form given in the last footnote) is the best possible.[9]

Arguably, the most beautiful and useful result obtained by Riemann in [Rie1] is the so-called Riemann 'explicit formula', connecting $\Pi(x)$ (or related counting

---

[7] Referring, in particular, to Riemann's groundbreaking *Habilitationschrift*—titled *On the Hypotheses at the Foundations of Geometry* and presented in 1854 to the University of Göttingen—Sir Arthur S. Eddington—the British astronomer whose observation of the 1919 total eclipse of the Sun first confirmed the bending of light rays grazing a massive body (like the Sun), as predicted by Einstein's theory of general relativity—made the following statement (quoted in [Ac,p.19]): "*A geometer like Riemann might almost have foreseen the more important features of the actual world.*"

[8] either in the form (1.4.1) or in the following (improved) form conjectured by Gauss,

$$\Pi(x) = Li(x)(1 + o(1))$$

as $x \to \infty$, where $Li(x) := \lim_{\epsilon \to 0^+} (\int_0^{1-\epsilon} + \int_{1+\epsilon}^{x}) \frac{1}{\log t} dt$ denotes the logarithmic integral

[9] Namely, for every $\delta > 0$, $\Pi(x) = Li(x) + O(x^{\frac{1}{2}+\delta})$ as $x \to +\infty$; see, e.g., [Pat,§1.8 and §5.8].

functions) and the zeros of the Riemann zeta function. Indeed, it expresses a deep relationship between the prime numbers and the critical (or nontrivial) zeros of ζ. (See, e.g., [Edw,Chapter3], [In], [Pat,Chapter3], [ParSha1,§2.5], [TeMeF,§2.4] and [Lap-vF2,p.4 and pp.75–76].) Riemann's formula is sometimes referred to as the Riemann–von Mangoldt explicit formula (see, e.g., [Lap-vF2,§4.5]) because a suitable version of it was later proved rigorously by von Mangoldt [vM1,2] in the mid-1890's. (See Equation (2.4.20) in Section 2.4.1 below for a classic version of Riemann's formula.) We note that such an explicit formula—along with its later generalizations to other parts of number theory—has recently been extended to the setting of fractal geometry in [Lap-vF1,2] in order to develop the theory of complex dimensions of fractal strings and to precisely describe the oscillations intrinsic to the geometry or the spectrum of fractals in terms of the underlying complex dimensions. (See [Lap-vF2], Chapter 4 and the relevant applications discussed in Chapters 5–9; see also [Lap-vF9] for further extensions and improvements.) Earlier, in [LapMa1,2], a geometric reformulation of the Riemann Hypothesis was obtained in terms of a natural inverse spectral problem for the vibrations of fractal strings. Rephrased in a more pictorial language, the work of [LapMa1,2] can be seen as demonstrating that the question (à la Mark Kac [Kac1]) *Can one hear the shape of a fractal drum*?—suitably interpreted as the aforementioned inverse problem, connecting the geometric and spectral oscillations of a fractal string—is intimately connected with and, in fact, equivalent to the Riemann Hypothesis. This characterization of the Riemann Hypothesis was extended and placed in a broader context in [Lap-vF2], especially in Chapter 7. In particular, the intuitive picture of the critical strip $0 \leq \text{Re}\, s \leq 1$ for $\zeta(s)$—suggested by the work in [LapPo1,2] (see especially [LapPo2,§4.4b], along with [Lap2,Figure 3.1 and §5] and [Lap3,§2.1,§2.2 and p.150]) and corroborated by the results of [LapMa1,2]—has been rigorously justified in [Lap-vF1,2]. (See [Lap-vF2,Figure 7.1,p.165] and the discussion surrounding it.)

In my opinion, the importance of the Riemann Hypothesis does not lie solely in the incredible multiplicity of its equivalent forms, but also in the cryptic message which it carries with it: one about the geometry of a landscape thus far inaccessible to us, the landscape underlying the prime numbers, and hence the integers. Once we will have found the clues needed to decode this message, we should be able to discover and unify large new areas of mathematics, lying at the confluence of arithmetic and geometry.[10]

1.5. Motivations, Objectives and Organization of This Book

At least from the physical point of view, our goal in the present book is more modest than the earlier discussion may have suggested. Indeed, we will not attempt to develop a geometry which models physical reality at scales where quantum gravity plays an essential role. Instead, we will propose a geometric and physical model that may help us to better understand aspects of number theory, particularly the set of prime numbers (or of integers) and the associated Riemann zeta function, along

[10] Along similar lines, one could perhaps consider the Riemann Hypothesis—together with the related information concerning the statistical distribution of the prime numbers (for example, that connecting the critical zeros of the Riemann zeta function and aspects of random matrix theory [Mon,Ber3–4,Od1–2,Gut2,RudSar,KatSar1–2,BerKe, KeSn1–2])—as a mathematical analogue of the recent COBE (and WMAP) observations regarding the extraordinary uniformity (and the tiny fluctuations) of the Penzias–Wilson cosmic background radiation ([Tri,PuGis] and [HuW,Stra]).

with their various generalizations for algebraic number fields and curves over finite fields which arise naturally in arithmetic geometry (see, for example, [ParSha1,2]).

This new model is motivated in part by several mathematical and physical sources, including the following ones:

(i) The theory of *fractal strings* [Lap1–4,LapPo1–3,LapMa1–2, HeLap1–2, Lap-vF1–5,HamLap] (to be viewed here as 'fractal membranes', or equivalently, as 'quantized fractal strings') and the corresponding theory of (fractal or arithmetic) complex dimensions recently developed in the author's research monograph joint with Machiel van Frankenhuysen and entitled *Fractal Geometry and Number Theory*: *Complex dimensions of fractal strings and zeros of zeta functions* [Lap-vF2]. (See also the new book [Lap-vF9], *Fractal Geometry, Complex Dimensions and Zeta Functions*: *Geometry and spectra of fractal strings*, where the theory of complex dimensions developed in [Lap-vF2] is much further expanded.)

(i′) More generally, the study of the vibrations of *fractal drums*, associated with Laplacians (or, more general elliptic differential operators) on open sets with fractal boundary or on suitable (self-similar) fractals themselves. (See, for instance, [Lap1–6], [LapFl], [LapPo1–3], [LapMa1–2], [HeLap1–2], [KiLap1–2], [LapPan], [LapNRG], [GriLap], [Lap-vF1–5], [DauLap] and the relevant references therein related to the so-called 'Weyl–Berry Conjecture' [Wey1–2,Ber1–2].) We note that fractal strings correspond to the one-dimensional case of 'drums with fractal boundary' but also have certain features in common with the latter situation of Laplacians on fractals.

(ii) String theory (from theoretical physics) and its striking dualities, especially the so-called '*T-duality*', a key symmetry not present in ordinary quantum mechanics which enables us, for example, to identify physically two circular spacetimes with reciprocal radii. (See, e.g., [Asp,EvaGia,GivePR,Gree,Polc1–4,Schw2 –4,Va1–2,Wit14,16–17].)

(iii) Noncommutative geometry and the recent attempts to connect it with conformal field theory and string theory. (See, especially, [FroGa,ChaFro,Cha1–2,LiSz1–2,FroGrRe1–2].)

(iv) Recent attempts to connect aspects of noncommutative geometry and fractal geometry from several points of view. (See, especially, [ConSul], [Con6, §IV.3]—particularly [Con6,§IV.3(ε)], motivated in part by [LapPo1–2]—as well as [Lap3,Part II], [Lap5], [Lap6] and [KiLap2].)

(v) The intriguing work of Deninger [Den1–7] on a possible cohomological interpretation of analytic number theory, as well as on the Extended Weil Conjectures and, in particular, on the (Extended) Riemann Hypothesis.[11]

To avoid any possible misunderstanding, we note that because we will consider here the vibrations of fractal membranes rather than of fractal strings, the roles played by the Riemann zeta function and our proposed approach to the Riemann Hypothesis will differ significantly from their respective counterparts in the previous work of the author and of his collaborators, Carl Pomerance, Helmut Maier, Christina He and Machiel van Frankenhuysen [Lap1-4,LapPo1–3,LapMa1–2,HeLap1

[11] See, e.g., Appendix B to the present work, especially §B.2 and §B.3, for a brief discussion of the classic Weil Conjectures [Wei5] (and Theorem [Wei1–4], in the case of curves over finite fields), along with some of their motivations.

–2,Lap-vF1–5,9]. Nevertheless, the concepts, techniques and results of this earlier theory will serve as an important motivation and a useful guide in a variety of ways.

The rest of this book is organized as follows:

In Chapter 2, we discuss the simple but important model of (closed) string theory on a circle (or, more generally, on a finite-dimensional torus). We do so both from the standard physical point of view (in Section 2.2) and—following the work of Fröhlich and Gawędzki [FroGa], pursued by Lizzi and Szabo in [LiSz1,2]—from the point of view of noncommutative geometry (in Section 2.3). T-duality is presented from each perspective in Section 2.2 and Section 2.3, respectively. Recall that this duality identifies the physics of string theory on two circles of reciprocal radii (see §2.2.2). More generally, in higher dimension, T-duality identifies the physics of two toroidal spacetimes associated with a pair of mutually dual lattices (see Remark 2.2.2).

In Section 2.4, we suggest that in this context, the functional equation of the Riemann zeta function $\zeta = \zeta(s)$ is a natural counterpart of T-duality for string theory on a circle (or, more generally, on a fractal membrane, in the sense of Chapter 3), while the Riemann Hypothesis may be connected, in particular, to the existence of a fundamental (or minimum) length in string theory, itself a consequence of T-duality. (We point out to the interested reader that in the first part of Section 2.4, we review some of the basic properties of $\zeta(s)$—and of other number theoretic zeta functions—which are used throughout much of this work; see Section 2.4.1 .)

In Chapter 3, we then propose an extension of this model to string theory on an infinite dimensional (adelic) torus, or on a Riemann surface with infinite genus. This yields a geometric model of the vibrations of a fractal membrane, viewed as a multiplicative (or quantized) analogue of a fractal string.[12] For a suitable choice of data—directly expressed in terms of the sequence of prime numbers—the quantum partition function of such a model then coincides with the Riemann zeta function $\zeta(s)$. Thus, by analogy with statistical physics [YaLe,LeYa,Jul1–2], the complex zeros (and the pole) of $\zeta(s)$ may be interpreted as corresponding to phase transitions. We therefore obtain an alternate mathematical answer to Bernard Julia's question raised in [Jul1,2], apparently rather different from that provided earlier by Bost and Connes in [BosCon1,2] (see also the exposition in [Con6,§V.11]). (Recall that Julia's problem consists in finding a natural mathematical model for a quantum statistical system, called a 'Riemann gas', whose partition function is equal to the Riemann zeta function.) We point out, however, that our primary motivations and objectives in developing the theory of fractal membranes are much broader and more ambitious than in the latter work, as will be clear in Chapters 4 and 5 (especially, Sections 5.4 and 5.5).

More specifically, after having recalled in Section 3.1 some basic facts concerning the theory of fractal strings (e.g., [Lap3,LapPo2–3,LapMa2,HeLap2,Lap-vF2,9]) and of their associated complex dimensions [Lap-vF2,9], we introduce the new notion of a *fractal membrane* (in Section 3.2), along with its self-similar counterpart, called a *self-similar membrane* (in Section 3.3). This enables us, in particular,

[12] Alternatively, rather than an (adelic) infinite torus, a fractal membrane can be thought of as an adelic Hilbert cube, with opposite faces identified. By 'adelic', in this context, we mean physically that each normal mode of vibration of such an object involves only finitely many circles (or faces).

to provide a mathematical model[13] of many arithmetic geometries and to obtain a natural interpretation of a standard Euler product of the associated zeta function—defined as the (spectral) partition function $Z(s)$ of the corresponding fractal membrane. For example, the special case of the so-called *prime membrane* yields the classic Riemann zeta function: $Z(s) = \zeta(s)$, as discussed in the previous paragraph. Further, our constructions and results can be easily adapted to other 'prime membranes', associated with arbitrary algebraic number fields or with curves (or higher-dimensional varieties) over finite fields. In that case, $Z(s)$ coincides with the zeta function of the field or the zeta function of the curve, respectively. (See §3.2.1 and Example 3.2.14, along with §2.4.1 and §B.1 in Appendix B.) For a general fractal membrane, we point out that the role of the 'primes' is played by the radii lengths of the circles of the infinite dimensional torus associated with the membrane. Our main result in Section 3.2.2 can then be interpreted as stating that the partition function $Z(s)$ of the membrane coincides with the corresponding Beurling zeta function [Beu1]. (See Theorem 3.2.8 and the comments following it.) As was mentioned earlier, however, our primary goals and motivations in introducing the notion of a fractal membrane and developing its theory go well beyond the consideration of this particular problem. (See Chapters 4 and 5.)

In Section 3.3, we show that the partition function of a self-similar (rather than fractal) membrane is no longer given by a standard Euler product but instead coincides with the geometric zeta function of a self-similar string with infinitely many scaling ratios, which now play the role of the generalized primes. (This naturally extends earlier results in [Lap-vF1,2] obtained for standard self-similar strings with finitely many scaling ratios; see §3.1, especially Example 3.1.2.) In the process, we also develop and significantly deepen the analogy between arithmetic and self-similar geometries pointed out in earlier work of the author and his collaborators, particularly in [Lap3] and [Lap-vF2]. This analogy is used throughout much of the rest of this book in order to transfer concepts or results from one subject to the other.

We mention that near the end of Chapter 3 (more specifically, in Section 3.3.1), we also show that the partition function of a fractal membrane coincides with the (appropriately weighted) dynamical (or Ruelle) zeta function of a suitable 'suspended flow' (introduced in passing in [Lap-vF3]). This yields, in particular, a dynamical interpretation of the Euler product expansion of the partition function—or, equivalently, of the Beurling zeta function associated with the underlying generalized primes—in terms of the primitive (or 'prime') orbits of the flow. Accordingly, this flow may be called a '*Riemann–Beurling flow*' because the weights (or 'lengths') of its primitive orbits coincide with the logarithms of the underlying generalized primes of the membrane. Furthermore, we obtain the analogue of these results for self-similar flows (in the sense of [Lap-vF3,9]). In particular, we show that the dynamical zeta function of a self-similar flow coincides with the partition function of the associated self-similar membrane. We thereby extend to the case of infinitely many scaling ratios the dynamical interpretation of the geometric zeta function of a self-similar string that was obtained in [Lap-vF3] and [Lap-vF9,Chapter 7]. These new results complete the aforementioned analogy between fractal and self-similar

[13] This was recently made rigorous in a joint work in preparation with Ryszard Nest [LapNe1], where fractal membranes are shown to be, in a suitable sense, the second quantization of fractal strings. (See §4.2 for a brief account of these results.)

membranes. They may also be potentially very useful in future work exploring the possible spectral and cohomological interpretation of the dynamical complex dimensions which is conjectured to exist in the latter part of Chapter 4 (see §4.4).

In Chapter 4, entitled *Noncommutative models of fractal strings*: *fractal membranes and beyond*, we discuss increasingly rich and noncommutative models of fractal strings and membranes. In particular, in Section 4.2, we provide a noncommutative geometric and operator algebraic (as well as quantum field theoretic) model of fractal membranes. (In this case, the underlying algebras of 'quantum observables' is noncommutative.) More specifically, we briefly discuss rigorous joint work in preparation with Ryszard Nest [LapNe1] in which we show that fractal membranes (in the sense of Section 3.2) can be precisely defined and are the second quantization of fractal strings, corresponding to a suitable choice of quantum statistics—namely, Fermi–Dirac statistics in the first construction of fractal membranes presented in Section 4.2, and Bose–Einstein statistics in the second construction, given in Section 4.2.1. Analogously, self-similar membranes (in the sense of Section 3.3) are the second quantized version of fractal strings, associated this time with the choice of Gibbs–Boltzmann statistics. In short, in agreement with the author's original intuition explained in Chapter 3, it follows from [LapNe1] that fractal membranes (along with their self-similar counterparts) are truly *quantized fractal strings*, but now in a very precise mathematical sense.

A significant advantage of the aforementioned second construction (see §4.2.1) is that it enables one to define a fractal membrane as a true noncommutative geometric space (in Connes' sense, as discussed earlier in Section 1.2). Such a space is given by a suitable 'spectral triple' $(\mathcal{A}, \mathcal{H}, D)$, where $\mathcal{A}$ is a noncommutative C^*-algebra represented on a complex Hilbert space $\mathcal{H}$, and D is an unbounded, self-adjoint operator on $\mathcal{H}$ viewed as the 'Dirac operator' on the underlying noncommutative space. (See, e.g., [Con6].) Here, $\mathcal{A}$ plays the role of the 'algebra of quantum observables', the noncommutative 'algebra of coordinates' or the algebra of 'Lipschitz functions' on the underlying noncommutative space. Furthermore, the Hilbert space $\mathcal{H}$ can be thought of as a suitable 'Fock space' on which the 'Dirac-type operator' D acts. Additional desirable properties are satisfied by this spectral triple, as is explained in Section 4.2.1 and [LapNe1]. A new insight provided by this construction (from [LapNe1]) is that once a given fractal string has been 'quantized', its endpoints are no longer fixed in the real line but are instead free to move (or 'float') within a (holomorphic) disc in the complex plane. In hindsight, this is in some sense analogous to 'D-branes' [Polc3,4] in nonperturbative string theory and M-theory. Therefore, from this perspective, fractal membranes can perhaps be viewed as '*fractal D-branes*'.

In Section 4.3, we investigate an even richer physical, algebraic and noncommutative geometric model of fractal membranes, inspired by our discussion in Chapters 2 and 3 (especially in Sections 2.2, 2.3 and 3.2). More specifically, we consider a model of string theory in a fractal membrane, viewed alternatively as an adelic Riemann surface with infinite genus or an adelic infinite dimensional torus. We are thus led to introduce a vertex algebra and the corresponding Dirac operator(s) associated with each hole (or 'circle') in the Riemann surface with infinite genus (or the 'adelic torus')—see especially Section 2.3 from Chapter 2. In this context, it is good to keep in mind that heuristically, the radius of each circle of the infinite dimensional torus represents a (generalized) prime associated with the membrane.

(Recall that vertex algebras are algebraic structures used to describe quantum fields and their interactions in conformal field theory and in string theory. See the original references [BelaPZ], [Bor2], [FrenkLepM2]; see also, for instance, [Geb], [Kac-v] or [Polc3], and in the context of noncommutative geometry, [FroGa] and [LiSz2]. Moreover, see Appendix A of the present work.) Mathematically, this yields a sheaf of vertex algebras—or, more generally, of noncommutative spaces—providing an algebraic and geometric model for the quantum geometry underlying string theory in a fractal membrane. For a suitable choice of data, the resulting 'noncommutative stringy spacetime' may be an interesting model for exploring and trying to understand the geometry underlying the prime numbers, as well as the integers, which viewed multiplicatively, coincide with the frequencies or 'energy levels' of the membrane. Furthermore—since, as was mentioned above in our discussion of Chapter 3, our proposed construction can be extended to algebraic number fields as well as to curves (or higher-dimensional varieties) over finite fields, for example[14]—the resulting family (or 'moduli space') of quantum geometries may provide a natural model for Deninger's (heuristic) notion of an 'arithmetic site' [Den1,3,5–6,8]. (See §5.4.1.) As is discussed in several places in Chapter 5, this should be closely related to the notion of 'moduli space of fractal strings' introduced by the author in the early 1990's in order to provide a natural receptacle for many of the zeta functions arising in arithmetic and fractal geometry and to classify the various types of (one-dimensional) fractal geometries occurring in his theory of fractal strings and of their vibrations.

In Section 4.4, several conjectures are proposed—regarding fractal membranes and their (dynamical) complex dimensions[15]—that would yield new insights into the nature of the Riemann zeros and into the possible algebraic and geometric structures underlying the Riemann Hypothesis. In particular, we conjecture that a suitable spectral and cohomological interpretation of the dynamical complex dimensions of prime membranes—and notably, of the Riemann zeros—can be obtained in this context, by means of the associated sheaf of vertex algebras (or, more generally, of noncommutative spaces). This is partly inspired by Deninger's work on 'cohomological number theory' and the Extended Weil Conjectures (see §B.3 of Appendix B in conjunction with §4.4).

In closing this overview of Chapter 4, we mention that in Section 4.4.1, we very briefly discuss the possible connections between aspects of our work and Shai Haran's appealing approach to "*The mysteries of the real prime*" [Har2] and the Riemann Hypothesis.

In Chapter 5, we introduce the *moduli space of fractal strings* $\mathcal{M}_{fs}$—along with its 'quantization', the *moduli space of fractal membranes*, $\mathcal{M}_{fm}$—viewed as highly noncommutative (quotient) spaces, in the spirit of Connes' noncommutative geometry, and as a broad generalization of the set of all Penrose tilings (or of all quasiperiodic tilings of the plane). We analyze the zeta functions (or spectral partition functions) associated with these moduli spaces, and show that a significant advantage of the moduli space of fractal membranes $\mathcal{M}_{fm}$ over that of fractal strings $\mathcal{M}_{fs}$ is that both the poles *and the zeros* (rather than just the poles) of the corresponding zeta functions are natural geometric invariants—see Sections 5.1 and 5.2. In Section 5.3, we propose that since the moduli space of fractal strings (or its

[14] as is the case for the fractal membranes discussed in §3.2 and §4.2

[15] i.e., the poles and the zeros of the associated partition function

quantization, $\mathcal{M}_{fm}$) is a natural receptacle for zeta functions, it may be viewed as a possible mathematical model for (and a suitable extension of) Deninger's elusive *arithmetic site* [Den1,3,5–6,8]. In short, from our perspective, this arithmetic site can be thought of as a heuristic 'space' the 'points' of which are expected to be the zeta functions of number fields, function fields, along with more general arithmetic zeta functions. In Section 5.4.2, we begin by providing the necessary operator algebraic background material on the beautiful theory of *factors* of von Neumann algebras and the associated *modular theory*, which particularly enables one to consider corresponding noncommutative flows such as the '*modular flow*' which, in its various guises, plays a key role in the rest of this chapter. (See §§5.4.2a–c, along with §5.5.) Then, in the latter part of Section 5.4.2 (§5.4.2d and §5.4.2e), building upon results and ideas from the theory of operator algebras and noncommutative geometry [Con5,6] as well as aspects of Connes' recent noncommutative geometric approach to the Riemann Hypothesis (as developed in [Con10] and announced in [Con9]), we state a conjecture about the nature of $\mathcal{M}_{fm}$ and of the associated (continuous, noncommutative) *modular flow*. It would follow that, in some sense, the modular flow itself can be viewed as a suitable substitute for and extension of the so-called *Frobenius flow* on the arithmetic site, arguably one of the Holy Grails of modern arithmetic geometry.

In Section 5.5, we conclude the main part of this book by proposing a geometric and dynamical interpretation of the Riemann Hypothesis. This is formulated in terms of the modular flow on $\mathcal{M}_{fm}$—thought of as a (noncommutative) flow of zeta functions or, equivalently, as a flow of the (generalized) primes of the underlying membranes—and its counterpart on the Riemann sphere, a Hamiltonian flow on the space of associated 'complex dimensions' (i.e., of the corresponding poles and, especially, *zeros*). In particular, we conjecture that the '*self-duality*' of the functional equations satisfied by arithmetic zeta functions (such as the Riemann zeta function and other L-series) forces the flow of (critical) zeros to 'land' on the Equator of the Riemann sphere,[16] which in this picture corresponds to the critical line $\operatorname{Re} s = \frac{1}{2}$. (See, especially, Figures 1 and 2 near the beginning of §5.5.2, along with the cover of this book.) Accordingly, the truth of the Riemann Hypothesis would be due to the intrinsic (dynamical) stability of 'arithmetic geometries' or 'self-dual geometries' (as forming the 'arithmetic site') within the moduli space of fractal membranes. Therefore, our proposed approach would not only explain why the Riemann Hypothesis must be true but also provide a new geometric and dynamical framework within which to attempt to prove it.

More precisely, we conjecture that along the orbits of the flow of fractal membranes (on the 'effective part' of $\mathcal{M}_{fm}$), the corresponding generalized fractal geometries (viewed as noncommutative spaces) are continuously deformed (i.e., 'converge') to arithmetic geometries.[17] This implies that along a given orbit, the zeta functions (i.e., spectral partition functions) of the fractal membranes converge to the arithmetic zeta function associated with the limiting 'arithmetic geometry'.

---

[16] Recall that via stereographic projection, the Riemann sphere—defined as the complex plane completed by a point at infinity—can be identified with S^2, the unit sphere of the 3-dimensional Euclidean space $\mathbb{R}^3$.

[17] Arithmetic geometries—which form the 'core' of $\mathcal{M}_{fm}$—are thus the 'stable attractive fixed points' of the noncommutative flow. Further, they are viewed here as '*self-dual geometries*' (relative to a suitable counterpart of 'T-duality').

In particular, these zeta functions[18] become increasingly 'self-dual'. Furthermore, also by 'T-duality', it follows that their zeros are attracted by the Equator of the Riemann sphere (i.e., converge to some discrete subset of the 'critical line'), both within the lower and the upper hemispheres. Consequently, the critical zeros of the limiting arithmetic zeta function—towards which the aforementioned (orbit of) zeros must also converge—naturally satisfy the (Extended) Riemann Hypothesis. In other words, the 'core' of $\mathcal{M}_{fm}$—viewed as the site of arithmetic geometries and hence, as a possible realization in our context of Deninger's arithmetic site—is the *attractor* of the modular flow of fractal membranes. Similarly, the 'critical line' (i.e., the Equator) is (or rather, contains) the attractor of the corresponding flow of zeros on the Riemann sphere. In a nutshell, this is the essence of the conjectural picture which we are proposing near the end of Chapter 5. (See Sections 5.5.1 and 5.5.2, including Figures 1 and 2; see also Section 5.4.2, particularly §5.4.2d and §5.4.2e.)

We close this description of the main contents of the book by mentioning that in the last subsection of Chapter 5 (§5.5.3), we discuss some analogies and possible connections between our proposed approach to the Riemann Hypothesis via modular flows of zeta functions (and their associated noncommutative geometries) and several types of geometric, analytic or physical flows encountered in (or inspired by) various aspects of contemporary mathematics and physics. In particular, in §5.5.3b, we propose a model—called the 'KMS-flow for (generalized) Pólya–Hilbert operators'—of the modular flow of zeta functions and their zeros. This model is inspired in part by the operator algebraic approach to quantum statistical physics (see §5.4.2b, along with §5.5.3b) and by analogies with two different but complementary approaches (discussed in §5.5.3a) to the Schrödinger equation[19] with a highly singular potential. Namely, these are the approaches via analytic continuation (in 'mass' or in the 'diffusion constant') of a suitable Feynman path integral, or else via Wilson's renormalization flow (or group); see §5.5.3a. Furthermore, in §5.5.3c, we discuss possible analogies with the Ricci (–Hamilton) flow on (three-dimensional) manifolds, acting as a renormalization-type flow, as in the recent groundbreaking (and entirely independent) work of Perelman on Thurston's Geometrization Conjecture and, in particular, on the Poincaré Conjecture. Finally, whereas in §5.5.3c, the modular flow of zeta functions and the associated noncommutative geometries is suggested to be a suitable arithmetic and noncommutative counterpart of the Ricci flow, the corresponding flow of zeros is briefly viewed in §5.5.3d as a 'noncommutative, arithmetic and KP (or KdV) flow' acting as a geodesic flow on a certain noncommutative manifold. Although, admittedly, all of these flows arise in very different contexts, the analogies drawn in the various parts of Section 5.5.3 should provide a useful guide in future explorations of our proposed approach to the Riemann Hypothesis.

It may be helpful from the outset for the reader to be aware of the progression followed in this text and of the different nature of the various parts of this book (while also keeping in mind the intimate connections between them, as explained earlier). While the present introduction, Chapter 1, is aimed at a 'general' scientific audience (with a strong interest in mathematics and physics), Chapter 2 is more physics oriented and requires a certain familiarity with some of the basic aspects

[18] or rather, the 'generalized functional equations' which they satisfy

[19] viewed as a time-evolution equation

of quantum mechanics (and its modern incarnations). It has been written, however, with the mathematical reader in mind, and does not really require previous knowledge of string theory. The end of Section 2.3 (more specifically, the latter part of §2.3.2) is more technical and mathematical, and should perhaps be omitted upon a first reading. Suitable references to the relevant physics and mathematics literature (along with an appendix on the definition and properties of vertex algebras, Appendix A) are provided to facilitate the task of finding out more about the many fascinating subjects only touched upon in this chapter. It should be stressed that because string theory (or its recent nonperturbative extensions) is very far from having been experimentally verified, as was discussed towards the beginning of this introduction (see §1.3), our use of the term 'physical' in this context should be taken with a grain of salt. It is our point of view, however, that the beautiful mathematical structures revealed by string theory should have an important role to play in our future understanding of mathematical reality and in particular, of aspects of number theory and of arithmetic geometry.

The second part of this book, composed of Chapters 3 through 5, is of a more overtly mathematical nature than either Chapter 1 or 2. Chapter 3, for example, contains the statement of several definitions, theorems and proofs, more in the style of a traditional mathematical research monograph. Even then, however, some of the 'definitions' provided in Chapter 3 (in Sections 3.2 and 3.3) are only fully justified by rigorous joint research work in preparation [LapNe1] (motivated by the theory developed in this book and briefly discussed in Section 4.2). Large parts of Chapters 4 and 5 (with the exception of Sections 4.1 and 4.2) are certainly of a more speculative nature than most of Chapter 3. They build upon the material of Chapter 3 but also use or at least refer to a large amount of contemporary mathematics, as well as draw on the physical language and formalism introduced in Chapter 2. They also contain a number of conjectures, open problems and hypothetical statements suggested by our physical or mathematical discussion in Chapter 2 or 3. We hope that the reader will be able to adjust without too much difficulty to the different styles encountered in this book and to switch from one type of discourse to another—mathematical, physical, or speculative—sometimes within the same chapter or section, especially towards the end.

In order to facilitate this transition and make the book more readily accessible to a broader audience, we have included some relevant background material at various points in the text or in the appendices. See, in particular, Section 2.4.1 (on the Riemann and other arithmetic zeta functions), Section 3.1 (on fractal strings and their complex dimensions), Section 5.4.2 (which includes a review of modular theory and noncommutative flows on von Neumann algebras), as well as Appendix A (on vertex algebras) and Appendix B (on the classical Weil Conjectures and the Riemann Hypothesis for varieties over finite fields).

Vertex algebras provide an elegant algebraic language to describe the quantum interactions between point-particles or strings in conformal field theories (CFT's) or string theories, respectively, while the Weil Conjectures (for 'finite geometries') have served as a useful guide in the search for an appropriate strategy to tackle the original Riemann Hypothesis (within the context of a conjectural 'arithmetic geometry' associated with the Riemann zeta function, say) and its various extensions. The Weil Conjectures also partly motivate aspects of our discussion near the end of Chapters 2, 4 and 5.

Moreover, Appendix C gives a precise statement and proof of the general Poisson Summation Formula (PSF, in short) for a pair of dual lattices, along with some of its consequences. This formula plays a key role both in the physical and in the arithmetic situations discussed in Chapter 2. It also partly motivates several statements (or conjectures) made in Chapters 4 and 5. We note that the second part of Appendix C reviews aspects of the theory of modular forms and their associated L-series, whose functional equations are established by using the Poisson Summation Formula, and which are central to much of number theory and its various applications to other areas of mathematics and physics.[20]

Appendix D is devoted to a discussion of some of the most relevant analytic properties of Beurling zeta functions associated with systems of generalized primes (g-primes, in short). These zeta functions and the corresponding g-prime systems play an important role in Chapter 3 (especially, Section 3.2) and in parts of Chapters 4 and 5. Recall from our earlier discussion (and §3.2) that the spectral zeta function of a fractal membrane coincides with the Beurling zeta function of a g-prime system naturally associated with the membrane.

Furthermore, Appendix E on the 'Selberg Class of zeta functions' gives an overview of the basic properties of this class of arithmetic-like meromorphic functions. Some of these properties are already established, while others are merely conjectured at this point. The relevance of the Selberg Class to our work stems from our expectation that the notion of a fractal membrane and the corresponding moduli space of fractal membranes introduced in Chapters 3 and 5, respectively, can naturally be extended to include this family of meromorphic functions as associated spectral partition functions (or 'zeta functions').

Finally, in Appendix F, we give a more detailed and mathematical description of the noncommutative space of Penrose tilings considered in Section 5.1, via the notion of 'groupoid C^*-algebra' associated with the underlying singular (and, in particular, non-Hausdorff) quotient space. We also discuss extensions of this construction that can be used to associate suitable noncommutative spaces to quasicrystals (and to corresponding nonperiodic tilings). In the process, we review at some length several notions of mathematical quasicrystals and related concepts. The material and the constructions provided in this appendix should play an important role in the formalization of the notion of 'generalized fractal membrane' (viewed heuristically as some kind of generalized quasicrystal) and of the associated moduli spaces that are discussed or alluded to in Chapter 5 (especially, Sections 5.4 and 5.5), Appendix E and earlier places in the book.

For an introduction to the origins of the theory developed in this book, see [Lap8], the text from which the project of this book emerged, the content of which is based on the author's ideas and intuitions extending over many years.

It should be made clear to the reader that the research program outlined in this book is still at an early stage and that some of the mathematics involved or implied is rather formidable or even not yet formulated in a precise manner. Nevertheless, we hope that the ideas and models that we propose here can be suitably modified and/or extended in order to build a useful bridge between noncommutative, string,

[20]We note that modular forms—along with their higher-degree counterparts, automorphic forms, briefly considered in §E.4 of Appendix E—play an important role in our discussion of the 'arithmetic site', viewed as the 'core' of the moduli space of 'generalized fractal membranes'. (See esp. §5.4.2e.)

arithmetic and fractal geometry. We also hope that readers will be motivated by this book to further investigate the mysterious and elusive geometry underlying the prime numbers (thereby, the integers) and, of course, the Riemann zeros.