Physics of the Riemann Hypothesis

Dániel Schumayer and David A. W. Hutchinson

Jack Dodd Centre for Quantum Technology,
Department of Physics,
University of Otago, Dunedin,
New Zealand

(Dated:)

Physicists become acquainted with special functions early in their studies. Consider our perennial model, the harmonic oscillator, for which we need Hermite functions, or the Laguerre functions in quantum mechanics. Here we choose a particular number theoretical function, the Riemann zeta function and examine its influence in the realm of physics and also how physics may be suggestive for the resolution of one of mathematics’ most famous unconfirmed conjectures, the Riemann Hypothesis. Does physics hold an essential key to the solution for this more than hundred-year-old problem? In this work we examine numerous models from different branches of physics, from classical mechanics to statistical physics, where this function plays an integral role. We also see how this function is related to quantum chaos and how its pole-structure encodes when particles can undergo Bose-Einstein condensation at low temperature. Throughout these examinations we highlight how physics can perhaps shed light on the Riemann Hypothesis. Naturally, our aim could not be to be comprehensive, rather we focus on the major models and aim to give an informed starting point for the interested Reader.

CONTENTS

I. Introduction 1
II. Historical background and ‘Mathematical necessities’ 1
III. Connections to physics 4
   A. Classical mechanics 4
   B. Quantum mechanics 7
      1. Scattering state models 10
      2. Bound state models 13
   C. Nuclear physics 14
   D. Condensed matter physics 16
   E. Statistical physics 18
IV. Conclusion 23
Acknowledgments 24
References 24

I. INTRODUCTION

‘Can you do Addition?’ the White Queen asked.
‘What’s one and one and one and one and one and one and one and one and one and one and one and one and one and one and one and one and one and one and one?’
‘I don’t know,’ said Alice. ‘I lost count.’
(Lewis Carroll - Through the Looking Glass)

Counting, in the broadest sense, is probably the oldest mathematical activity and not even uniquely ours. Even animals can distinguish one, two and three, maybe just by recognising a pattern, but only humans have developed an abstract language, mathematics or more specifically number theory, which accurately describes the properties of numbers.

In the following we will focus on the border between physics and number theory, and more precisely, how the Riemann-zeta function, \( \zeta(s) \), appears in quite different areas of physics. This review does not intend to be comprehensive, rather would like to offer a panoramic view and give a feeling as to why many physicist find beauty in the structure of this seemingly random function and what one might learn from it. We collect examples from diverse realms of physics, from classical mechanics to condensed matter physics, where the Riemann-zeta function or its ‘descendants’ play a significant role. Due to space limitations we do not aspire to be mathematically precise in our derivations, but we give physical arguments to support results, and also direct the Reader to relevant sources.

II. HISTORICAL BACKGROUND AND ‘MATHEMATICAL NECESSITIES’

God invented the integers;
all else is the work of man.
(Leopold Kronecker)

Natural numbers form the basis of our arithmetic, with various operations defined among these numbers. All of us learn to use four basic operations: addition, subtraction, multiplication and division. The latter, division, hides one of the most enigmatic internal structures of the set of the natural numbers, namely that there are special
numbers, the primes, among the natural numbers which cannot be divided by any other natural number, other than unity and themselves, without a remainder. Euclid of Alexandria proved that there are infinitely many such numbers. Later, Eratosthenes of Cyrene gave a theoretical algorithm, a sieve, for finding these primes amongst the natural numbers. Despite all efforts in the last two thousand years, the efficient determination as to whether a given number is prime or not still proves a remarkable challenge.

It is not hard to understand why the distribution of primes could captivate the imagination of many mathematicians and physicists. These numbers seem to obey two contradictory principles. Firstly, they seem to appear randomly among composite numbers, but secondly they also appear to obey strict rules governing their distribution.

Apart from Euclid’s, numerous proofs exist for the infinitude of the prime numbers (Ribenboim, 1991). Euler, at the early age of 30, proved a stronger statement (Euler, 1737),

$$\sum_{p \text{ prime}} \frac{1}{p} = \infty. \quad (1)$$

This formula clearly proves Euclid’s statement but it also demonstrates the frequent occurrence of prime numbers amongst composite numbers. A natural continuation of his work was the analysis of the arithmetic properties of the series, $\sum n^{-k}$. Substituting $k = 1$ into this expression we recover the well-known, divergent harmonic series. Conversely, if $k > 1$ the summation converges. Euler also showed (Euler, 1737) – using the fundamental theorem of arithmetic – that this series can be written as an infinite product over the prime numbers, $p$, such that

$$\zeta(k) = \sum_{n=1}^{\infty} \frac{1}{n^k} = \prod_{p} \left(1 - \frac{1}{p^k}\right)^{-1}. \quad (2)$$

One may interpret through this relationship that the prime numbers construct the $\zeta(k)$ function. Since $p$ denotes a prime number and $k > 1$, none of the factors in this product can be zero. Therefore we can conclude that $\zeta(k)$ does not have any zeros if $k > 1$.

Bernhard Riemann, who was the first to apply the tools of complex analysis to this function, proved that the function defined by the infinite summation (Riemann, 1859)

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad (3)$$

can be analytically continued over the complex $s$ plane, except for $s = 1$. This analytic continuation of the function is called the Riemann-zeta function. Here we follow the traditional notation, with $s$ denoting a complex number, $s = \sigma + it$, where $\sigma$ and $t$ are real numbers and $i$ is the usual imaginary unit.

Riemann also derived a functional equation, containing the $\zeta(s)$ function, which is valid for all complex $s$ and exhibits mirror symmetry around the $\sigma = 1/2$ vertical line, called the critical line, such that

$$\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \pi^{-\frac{1-s}{2}} \Gamma\left(\frac{1-s}{2}\right) \zeta(1-s) \quad (4)$$

One should note that the zeta function stands on both sides, on the left hand side with argument $s$, while on the right hand side with $(1 - s)$. This relationship between $\zeta(s)$ and $\zeta(1 - s)$ provides some insight regarding the location of the zeros of this function. Let us examine the half-line for which $\sigma < 0$, and $t = 0$. The products on either side can be zero if at least one of the factors is zero. On the right hand side of (4) all the pre-factors of the zeta function are non-negative and do not have any zeros. On the other side, however, the $\Gamma(\sigma/2)$ function has simple poles at all even negative integers. The equation can hold only if $\zeta(\sigma)$ has simple zeros at the same locations. These zeros are called trivial, because their locations are inherited from the $\Gamma$ function. The same argument also shows that all other zeros of the $\zeta(s)$ function have to lie in the $0 \leq \sigma \leq 1$ region, called the critical strip. The zeros located in this strip are the non-trivial zeros of the Riemann-zeta function. It can also be shown that the non-trivial zeros $\rho$ are arranged symmetrically, both in respect of the critical line and the $t = 0$ axis. Figure 1 depicts the pole and zero structure of $\zeta(s)$ on the complex

![Figure 1](image-url)

**FIG. 1** The ‘anatomy’ of the Riemann-zeta function on the complex $s$ plane. The black dots (●) represent the zeros of ζ(s), including possible zeros which do not lie on the critical line.
s plane including the possible zeros off the critical line.

So far the statements about the zeros of $\zeta(s)$ and their locations on the complex plain were simple. However the distribution of the non-trivial zeros holds one of the most intriguing and enigmatic mathematical mysteries of the last century and a half. It is embarrassingly easy to pose Riemann’s conjecture: all non-trivial zeros of $\zeta(s)$ have the form $\rho = 1/2 + it$, where $t$ is a real number. In other words all non-trivial zeros lie on the critical line. In 1900 Hilbert nominated the Riemann Hypothesis as the eighth problem on his famous list of compelling problems in mathematics (Hilbert, 1902). Since then not just professional mathematicians but mathematical soldiers of fortune tried, and still try, to verify its validity. The stakes are high. Whoever proves or disproves this hypothesis engraves his name in the tablets of the history of mathematics, and may also receive one million dollars from the Clay Mathematics Institute.

During the past century, the Riemann Hypothesis has been recast into many equivalent mathematical statements. A few of them are purely number theoretical in origin, such as the Mertens conjecture, which we will later discuss in the context of a special Brownian motion, but other redefinitions are very much cross-disciplinary. A more advanced mathematical introduction to the history of the Riemann Hypothesis and its equivalent statements can be found in an excellent monograph and compendium (Borwein et al., 2008) which is readable not just at the expert, but also the undergraduate level.

The distribution of the $\zeta(s)$ zeros, with real part equal to 1/2, has thus attracted significant interest. One of mathematics’ giants has proven that infinitely many zeros do lie on the critical line (Hardy, 1914), however Riemann’s conjecture is much stronger, requiring all the zeros to be on the critical line. In 1942 Selberg proved

$$N_0(T) > C T \ln(T) \quad (C > 0 \text{ and } T \geq T_0) \quad (5)$$

i.e. the number of zeros of the form $s = \frac{1}{2} + it$ ($0 \leq t \leq T$), denoted by $N_0(T)$, grows as $T \ln(T)$ at least for large $T$.

Three decades later, in 1974, Levinson showed that at least one third of the non-trivial zeros are on the critical line (Levinson, 1974) which was later incrementally improved to two fifths (Conrey, 1989). This small step over a period of twenty years is indicative of the difficulty of the Riemann Hypothesis.

Let us return to the linkage between the $\zeta(s)$ zeros and prime numbers. Equation (2) clearly shows the strong connection between the $\zeta(s)$ function and the prime numbers. This relationship can be made even more explicit if one examines how the number of primes below a given threshold behaves as this threshold is increased. Based on empirical evidence, many mathematicians, e.g. Legendre, Gauss, Chebyshev (Dickson, 2005), have conjectured that the prime counting function, $\pi(x) = |\{p \mid p \text{ is prime and } p \leq x\}|$, asymptotically behaves as the logarithmic integral Li$(x)$. This conjecture is known nowadays as the Prime Number Theorem after Hadamard (Hadamard, 1896) and de la Vallée-Poussin (de la Vallée-Poussin, 1896) independently gave rigorous proofs of this statement. Interestingly, this theorem has a geometrical interpretation: the Prime Number Theorem is equivalent to the assertion that no zeros of $\zeta(s)$ lie on the $\sigma = 1$ boundary of the critical strip.

Riemann published (Riemann, 1859), although Mangoldt provided the rigorous proof (von Mangoldt, 1895), the following explicit formula for the prime-counting function $\pi(x)$

$$\pi(x) = \sum_{n=1}^{\infty} \frac{\mu(n)}{n} J\left(x^{1/n}\right) \quad (6)$$

where

$$J(x) = \text{Li}(x) - \lim_{T \to \infty} \left[ \sum_{\rho \leq T} \text{Ei}(\rho \log(x)) \right] + \int_{x}^{\infty} \frac{dt}{t^2 - 1} \frac{1}{t \log(t)} - \log(2).$$

Here $\mu(n)$ is the M"obius function, $\rho$ denotes the non-trivial zeros of the Riemann $\zeta(s)$ function, and Li$(x)$ and Ei$(x)$ stand for the logarithmic and exponential integral$^1$, respectively. Therefore, whoever knows the distribution of the non-trivial zeros of $\zeta(s)$, will also know the distribution of the prime numbers.

Selecting only the first terms of the summand in equation (6) reproduces exactly the Prime Number Theorem, i.e.

$$\pi(x) \approx \text{Li}(x) \approx \frac{x}{\ln(x)} \quad (7)$$

This observation may lead us to conclude that Li$(x)$ gives the main contribution to $\pi(x)$ while the other terms represent corrections, similar to a perturbative calculation in physics – an analogy to which we will return. Figure 2 depicts the prime counting function, $\pi(x)$ and its various approximations. One may notice that the leading

---

1 See http://www.claymath.org/millennium/Riemann_Hypothesis

$^2$ The Möbius function is defined as follows: $\mu(1) = 1, \mu(n) = 0$ if $n$ has a square divisor, and $\mu(p_1p_2\cdots p_k) = (-1)^k$ if all $p_i$s are different. Thus $\mu(2) = -1$ and $\mu(12) = 0$, and $\mu(21) = 1$.

$^3$ The notation for the logarithmic integral is ambiguous in the literature. There are two definitions

$$I_1(x) = \int_{0}^{x} \frac{dt}{\ln(t)} \quad \text{and} \quad I_2(x) = \int_{2}^{x} \frac{dt}{\ln(t)}$$

where $I_1$ is interpreted as a Cauchy principal value. These integrals differ only by a constant number. Depending on the book the Reader may consult, either $I_1(x)$ or $I_2(x)$ is denoted with Li$(x)$. Here, we prefer the former.
term, \( \text{Li}(x) \), captures the tendency of \( \pi(x) \) well and the appearance of the oscillations clearly show how the zeros \( \rho_n \) influence and refine the agreement. As \( x \to \infty \) the curves of \( \text{Li}(x) \) and \( \pi(x) \) will practically coincide on a similar plot.

![Graph showing the approximation for the prime counting function](image)

**FIG. 2** Figure depicts the approximation for the prime counting function, \( \pi(x) \) (dashed line), using only the first term, \( \text{Li}(x) \) (dash-dotted line), and using the first ten non-trivial pairs of zeros of the Riemann \( \zeta(s) \) function (solid line). In the inset we restricted the range to \([2, 10] \) and used the first 30 non-trivial zeros.

One may define a density for the complex, non-trivial Riemann-zeta zeros as

\[
d(N) = \sum_k \delta(N - \rho_k)
\]

where \( \delta \) is the Dirac-delta distribution. Following Sir Michael Berry (Berry, 1985) the spectral density can be separated into a smooth and an oscillatory part, \( d(N) = \bar{d}(T) + d_{osc}(T) \), as

\[
\bar{d}(T) = \frac{1}{2\pi} \ln \left( \frac{T}{2\pi} \right) + 1 - \frac{1}{2\pi} + O(T^{-1}) \quad (9a)
\]

\[
d_{osc}(T) = -\frac{1}{\pi} \sum_p \sum_{r=1}^{\infty} \frac{\ln(p) \cos(\pi r \ln(p))}{\sqrt{p^r}} \quad (9b)
\]

where the external summation of \( d_{osc}(T) \) runs over the prime numbers, \( p \). The oscillatory part, therefore, gives the fluctuations as individual contributions from each prime number \( p \) labeled by an integer \( r \) corresponding to the prime power \( p^r \). Based on the smooth density of Riemann-zeros one may derive the number of positive, non-trivial zeros up to a fixed value of \( T_0 \):

\[
N(\rho < T_0) = \int_0^{T_0} \bar{d}(T) \, dT = \frac{T_0}{2\pi} \ln \left( \frac{T_0}{2\pi} \right) - \frac{T_0}{2\pi} \quad (10)
\]

Changing variable to \( T = \ln(T_0/2\pi) \) and recasting our result using \( T \) we obtain

\[
N(T) \propto e^T \quad (11)
\]

i.e. the number of \( \zeta(s) \) zeros below \( T \) increases exponentially. Although at this point this change of variable seems somewhat arbitrary, we will see later that it further strengthens the similarity between the zeros of \( \zeta(s) \) and the periodic orbits of a chaotic system, where the number of periodic orbits also increases exponentially.

Finally, we note the fruitful and diverse area of extensions of the Riemann-zeta function. These generalised zeta-functions do also occur throughout physics, primarily in modern quantum field theories. This topic, however, is far beyond the scope of this short review and we can only suggest Elizalde’s monograph as an introduction and Lapidus’ book for a more authoritative study.

### III. CONNECTIONS TO PHYSICS

*The Riemann Hypothesis is a precise statement, and in one sense what it means is clear, but what it’s connected with, what it implies, where it comes from, can be very unobvious.*  
(Martin Huxley)

#### A. Classical mechanics

In this section we discuss those models of classical mechanics, such as billiards, which lead to the introduction of the notion of integrability and chaos. This development of ideas gave birth to a new paradigm, since it provided an insight into how the spectrum of quantised analogues of classical systems are connected to classical paths.

Classical mechanics, in its Lagrangian and Hamiltonian forms, is the exemplar for physics in the modern sense. The major theories, e.g. statistical mechanics, quantum mechanics, are first expressed in the language of analytical mechanics with the development traced to the Enlightenment. Although a few analytically solvable models, e.g. Kepler two-body problem, harmonic oscillator, gave confidence in the machinery of mechanics, it was soon realised that there are important cases, e.g. three-body problem, where one not just cannot solve the equations of motion analytically, but the motion is proven to be chaotic (Celletti and Perozzi, 2007). This behaviour is very peculiar and at first sight seems puzzling, since the governing equations are deterministic, yet the actual motion seems to behave randomly. The celestial relevance of this three-body problem was so fundamental and enticing that King Oscar II of Sweden and Norway offered a prize for the person who could solve the following problem (Barrow-Green, 1994)

For an arbitrary system of mass points which attract each other according to Newton’s law, assuming that no two points ever collide, give

[Elizalde, 1995]
[Lapidus, 2008]
the co-ordinates of the individual points for all time as a sum of a uniformly convergent series whose terms are made up of known functions.

Although this problem had not been solved, Poincaré was awarded this illustrious prize for his impressive contribution. His work revolutionised the analysis of such chaotically behaving systems, although one had to wait nearly a hundred years for this revolution to really happen.

In classical mechanics we distinguish a special class of systems, the integrable dynamical system, which possess as many independent integrals of motion, \( I_n \), (action variables) as degrees of freedom, \( N \). For these systems the Hamiltonian can be expressed as a function of these action variables, namely \( \mathcal{H} = \mathcal{H}(I_1, \ldots, I_N) \), and the equations of motion (\( n = 0, 1, \ldots, N \))

\[
\frac{d\varphi_n}{dt} = -\frac{\partial \mathcal{H}}{\partial I_n} \quad \text{and} \quad \frac{dI_n}{dt} = \frac{\partial \mathcal{H}}{\partial \varphi_n} \tag{12}
\]

are easy to solve: \( I_n = \text{constant} \) and \( \varphi_n = \varphi_{n,0} + \omega_n t \).

A theorem of topology then guarantees that these \( N \) constants of motion, provided they are independent of each other, define an \( N \) dimensional torus and each trajectory with constant energy lies on that torus. Therefore, as a specific case, the dynamics described by a one-dimensional time-independent Hamiltonian is necessarily integrable. In order to consider chaotic dynamics one has to either introduce a time-dependent Hamiltonian or increase the degrees of freedom to two or higher.

One of the ‘simplest’ generic models with two or more degrees of freedom is that of classical billiards. These are dynamical systems where a particle has constant energy and moves in a finite volume, which may contain impenetrable obstacles. Whenever the particle reaches the boundary it suffers specular reflection. Depending on the shape of the billiard, the motion can be integrable or chaotic. The analysis of a circular billiard (see Figure 3) is straightforward due to the rotational symmetry. The incident angle remains the same at each bounce and each impact can be calculated from the previous one by rotating the circle twice that angle. Therefore if the incident angle is a rational multiple of \( \pi \), i.e. \( m\pi/n \), the trajectory is periodic with period \( n \) and therefore finite, otherwise it is infinite. In this latter case the points where the ball hits the wall will be uniformly distributed along the circumference of the circle. It was also proven by Jacobi that in the latter case every interval of the circle contains points of the trajectory.

Before we step beyond billiards and generalise the idea of periodic orbits, the origin of trace formulae, let us make a short detour around a recent result (Bunimovich and Dettmann, 2003) regarding the circular billiard (see Figure 4). As we discussed, due to rotational symmetry, or in other words, the conservation of angular momentum, this billiard model is integrable and the trajectory is fully described by two angles, \( \beta \) and \( \psi \), the angle around the circumference measured from a pre-determined point and the incident angle of the trajectory at the boundary, respectively. With these variables the dynamics is governed by the mapping: \( (\beta, \psi) \rightarrow (\beta + \pi - 2\psi, \psi) \), where all angles are taken modulo \( 2\pi \) and the ball travels with unit velocity. The phase space of this system can be described by Birkhoff’s coordinates constructed from two angles: the arc-length coordinate \( q = \beta \) (measured in radians and modulo \( 2\pi \)), and the tangential momentum coordinate defined as \( p = \sin(\psi) \). By convenient normalisation, the arc-length of the billiard is unity and the velocity of the ball is also unity, the phase space is restricted to \( 0 \leq q < 2\pi \), and \(-1 < p < 1 \). This choice also introduces a natural unit time-step, the time elapsed between consecutive bounces, \( \Delta t = 2\cos(\psi) \). The movement of the ball can, therefore, be represented by a possibly infinite series of points inside this phase-space area. Despite the rather artificial appearance of this model, the electromagnetic field in optical or microwave cavities can be modeled by such billiards (Alt et al., 1998; Harayama et al., 2003; Nockel et al., 1997; Stockmann and Stein, 1990). Since these experimental billiards are not ideal, it is interesting to examine what happens to the dynamics of this system if we cut a small window(s) along the reflective boundary, thereby, naturally introducing dissipation or ‘leakage’. It is natural to ask: what is the probability, \( P(n) \), of a ball leaving the billiard after \( n \) bounces, what is the mean number of bounces, \( \langle n \rangle \), before the ball escapes, or similarly, what is the probability, \( P(t) \), that escape takes at least time \( t \).

For strongly chaotic billiards the latter probability decays exponentially, while for integrable billiards, such as the circular one, it softens to only power-law decay (Bauer and Bertsch, 1990) and can be qualitatively understood using a simple geometrical argument. The prob-

\[
\begin{align*}
\text{FIG. 3} & \quad \text{A circular billiard and a Bunimovich stadium, which is} \\
& \text{a rectangle smoothly joined by semi-circles. Two different} \\
& \text{types of trajectories, periodic orbits (1) and non-periodic} \\
& \text{trajectories (2) are also depicted.}
\end{align*}
\]
ability, \( p \), that the ball escapes in a bounce is proportional to the size of the gap to that of the boundary, \( p = \epsilon / L \). Moreover, the probability that the ball survives the first \((n - 1)\) bounces and escapes only at the \( n \)th bounce is \((1 - p)^{(n-1)p}\). Therefore the mean number of bounces occurring until escape is

\[
\langle n_{\text{escape}} \rangle = \sum_{k=1}^{\infty} k (1 - p)^{(k-1)p} = \frac{1}{p} \ln \frac{1}{\epsilon}.
\]

(13)

Let us now cut two (possibly overlapping) holes, with sizes \( \epsilon \), on the boundary and examine the non-escaping periodic orbits. Based on the geometrical argument used above, we expect the probability to be \( \sim 2/\epsilon \), if the two holes do not overlap. However, in systems where the trajectories do not diverge strongly, i.e. Liapunov exponent is close to zero, only a small fraction of the trajectories will eventually hit the opening on the boundary, and the mean escape time will be proportional to \( \epsilon \).

If the initial incident angle is taken to be \( \psi_{m,n} = \pi/2 - m\pi/n \), where \( m < n \) are integers and relative primes to each other, then the trajectory is closed and its period is \( n \). Let us now examine only those initial conditions for which the escape time is at least \( t \), or in other words, the number of bounces is at least \( N = [2\pi/\epsilon] \). To fulfill this requirement one might take the initial value of \( \psi = \psi_{m,n} + \eta \), where \( 0 \leq \eta < \epsilon \) and \( \beta \) can be restricted to the following range

\[
\beta^0 \in \left( \epsilon + \frac{n\eta}{\cos(\psi_{m,n})}; \beta \right) \bigcup \left( \beta + \epsilon + \frac{n\eta}{\cos(\psi_{m,n})}; \frac{2\pi}{n} \right).
\]

The prime indicates that angles are taken modulo \( 2\pi/n \). The probability can, therefore, be calculated if one sums up all possible values of \((m, n)\) pairs. This is the point where number theory enters into this physical problem; we have to guarantee that \( m \) and \( n \) are relative primes. Integrating over the permitted region of \( \beta^0 \) one may find

\[
\mathbf{P}(t, \epsilon, \theta) \sim \frac{1}{t} \sum_{n=1}^{N} n \mathcal{F}(n) \sum_{m} \left[ 1 - \cos \left( \frac{2m\pi}{n} \right) \right]
\]

(14)

where the exact form of \( \mathcal{F}(n) \) can be found in [Bunimovich and Detteman, 2003]. Surprisingly the sum over \( m \) can be explicitly determined. The first, unit term, simply counts how many numbers are relative prime to \( n \) and, therefore, it can be formally expressed using a special function of number theory; Euler’s totient function \( \phi \). The second term in equation (14) is also a special expression. If the summation were over all the integer numbers smaller than \( n \), one could connect it to the Fourier series. However, here one only uses those \( m \)’s which are relative primes to \( n \). Converting the cosine term to complex exponentials and using Ramanujan’s identity\(^5\) for the sum of exponentials, the contribution of the cosine term turns out to be another special function of number theory which we have already met, the Möbius function, \( \mu(n) \). Therefore, the probability of non-escaping orbits is

\[
\mathbf{P}_\infty = \lim_{t \to \infty} \left( t \mathbf{P}(t, \epsilon, \theta) \right) \sim \sum_{n=1}^{\infty} n [\phi(n) - \mu(n)] \mathcal{F}(n)
\]

(15)

The leading order behaviour of \( \mathbf{P}_\infty \) as a function of \( \epsilon \) can be determined by calculating its Mellin-transform

\[
\tilde{\mathbf{P}}(s) = \int_{0}^{\infty} \mathbf{P}_\infty(\epsilon, \theta) \epsilon^{s-1} \, d\epsilon
\]

(16)

and examining the residues of \( \tilde{\mathbf{P}}(s) \) on the complex \( s \)-plane. Bunimovich and Detteman showed that for the two-hole problem, where these holes are separated by \( 0^\circ \), \( 60^\circ \), \( 90^\circ \), \( 120^\circ \), \( 180^\circ \), the probability \( \tilde{\mathbf{P}}(s) \) is uniquely determined by the Riemann-zeta function, \( \zeta(s) \), by i.e. its pole and non-trivial zeros. The first corrections to the leading order term are given by the non-trivial zeros of \( \zeta(1 + s) \), which are of the order \( \sqrt{T} \ln(e)^{n-1} \), provided for all zeros of \( \zeta(s) \), \( \Re(s) = \sigma = \frac{1}{2} \) with multiplicity \( m \). The Riemann Hypothesis is then shown to be equivalent to different asymptotic estimates on the number of zeros [Titchmarsh and Heath-Brown, 2003]. Therefore, if the non-trivial zeros provide the second order corrections to the probability, it is instructive to examine the deviation of these probabilities experimentally from the leading-order geometric terms, namely

\[
\lim_{\epsilon \to 0} \lim_{t \to \infty} \left( e^{\delta - 1/2} \left[ \mathbf{P}_1(t) - \frac{2}{\epsilon} \right] \right) = 0
\]

(17a)

\[
\lim_{\epsilon \to 0} \lim_{t \to \infty} \left( e^{\delta - 1/2} \left[ \mathbf{P}_1(t) - 2t \mathbf{P}_2(t) \right] \right) = 0
\]

(17b)

where \( \mathbf{P}_1, \mathbf{P}_2 \) belong to the one- and two-hole problem, respectively. If it is (experimentally) found that for every \( \delta > 0 \) these equations are fulfilled, then it proves the validity of the asymptotic formulae, thus the validity of the Riemann Hypothesis. The numerical results by Bunimovich and Detteman does not contradict these

---

\(^4\) Euler’s totient function, \( \phi(n) \), gives the number of positive integers smaller than \( n \), which are relative prime to \( n \), e.g. for any prime number \( \phi(p) = p - 1 \), since all integers smaller than \( p \) are relative prime to \( p \).

\(^5\) Ramanujan’s sum is defined as

\[
c_n(m) = \sum_m e^{2 \pi i m/n}
\]

where the summation is over those values of \( m \), which are relative prime to \( n \). Using Möbius inversion for this sum one can prove that \( c_n(m) = \mu(n) \) [Hardy and Wright, 1961].
It can be shown rigorously for a generic classical chaotic system \cite{Cvitanovic1991}, that

\[
\text{Tr } (L(t'; r), r) = \sum_p T_p \sum_{r=1}^\infty \frac{\delta(t - rT_p)}{|\text{det}(1 - J_p)|}, \tag{19}
\]

where the first summation runs over the periodic orbits labeled by \( p \), while the second takes into account all repetitions, \( r \). \( J_p \) is the Jacobian matrix of \( F \) localised around the periodic orbit, also called the monodromy matrix.

Here we can make an important observation: although this equation looks cumbersome, it does relate the spectrum of the evolution operator to a global behaviour of periodic orbits. Therefore these two sets of abstract objects are intimately related to one another. The connection of this trace formula and its quantum mechanical counterpart to the Riemann zeta function will become clear in the next section.

\section{B. Quantum mechanics}

\textit{Below we expound the Pólya-Hilbert conjecture. We enumerate the one-dimensional Hamiltonians proposed for which the distribution of energy eigenvalues mimic the non-trivial zeros of the Riemann zeta function and analyse their relationship with the Gutzwiller trace formula. We also examine the possible symmetries of a ‘Riemann-operator’ since it partially encouraged the development of quantum mechanics with only CT or PT symmetry.}

In the dawn of the 20th century Bohr postulated a series of rules for describing the spectrum of the hydrogen atom well before the birth of Schrödinger’s and Heisenberg’s quantum mechanics. In these early days ‘quantisation’ meant to restrict the possible values of action variables of the classical system (Bohr-Sommerfeld, Wentzel-Kramers-Brillouin, etc.) and the rules worked well, up to an additive constant. However, this description cannot be satisfactory in general, since for the majority of classical systems the only constant of motion is the energy, and therefore a method of quantisation relying on the existence of action-angle variables could not be applied \cite{Einstein1917}. On the other hand, we know that classical mechanics works well for large systems, therefore quantum mechanics must give the same predictions for a large system as classical mechanics (Bohr’s correspondence principle). This unproven principle ties these two theories firmly together and the same principle inspired the use of the Riemann \( \zeta \) function in investigating the relationship of classical to quantum mechanics.

The basic question is: how can we quantise a classical mechanical system? Could we state anything about the spectrum of a quantum system, at least qualitatively, without solving the corresponding Schrödinger equation?

We cannot expect to be able to infer the complete spectrum of a generic system, but asking only for the average density of states may prove feasible. One can give a

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{billiard.png}
\caption{A circular billiard with small openings \cite{Bunimovich2003}.}
\end{figure}
crude, although remarkably precise, estimate: each quantum state occupies approximately $\hbar^2/2\pi$ phase-space, where $\hbar$ is the Planck constant divided by $2\pi$ and $f$ is the number of degrees of freedom. This result is rather general and the individual quantum systems differ only in the ‘fluctuations’ around this average. It turns out that the type of fluctuation depends on the behaviour of the classical counterpart; classically regular and chaotic systems are quite different. For example, a classical rectangular billiard is integrable (regular), whereas its quantum analogue exhibits a chaotic spectrum, with the spacing of the quantum levels, $s = \epsilon_n - \epsilon_{n-1}$, following an exponential distribution $P(s) \sim e^{-s}$ (see later in section III.C). However, a stadium billiard is classically chaotic, but the spectrum of its quantum counterpart is regular, meaning that $P(s)$ is small for small values of $s$ and sharply peaked at a finite value indicating a regular distribution of energy levels. We will see that interpreting the $\zeta(s)$ zeros as energy levels their distribution is breathtakingly similar to those of a quantum system’s. This has inspired physicists to examine whether one could associate a dynamical system with the Riemann zeta function.

The advantage of this approach would be that the huge number of $\zeta(s)$ zeros are known and quick numerical algorithms have also been developed to find further zeros, thus solving the Schrödinger equation for large energies would be unnecessary. The Riemann zeta function could play the same role in the examination of chaotic quantum systems as the harmonic oscillator does for integrable quantum systems. This is the point where the examination of the Riemann zeta function may help to understand physics or, vice versa, the physics may lead us to the solution of this so far intractable mathematical problem.

In order to establish a strong formal connection between a generic chaotic quantum system and the distribution of the Riemann $\zeta(s)$ zeros, we have to elucidate a new description of quantum systems, the Gutzwiller’s trace formula. This trace formula is the analogue of equations (19b) for physical systems.

Let us, therefore, return to a classically integrable system, for which the Hamiltonian $H$ can be given in terms of conserved quantities $H = \mathcal{H}(I_1, \ldots, I_N)$. Using Bohr’s semiclassical quantisation rules, these action variables take not arbitrary, but fixed values

$$I_k = \hbar \left( n_k + \frac{\mu_k}{4} \right), \quad (k = 1, 2, \ldots, N) \quad (20)$$

where the $\mu_k$ are integers and called Maslov indices (Arnol’d, 1997). The density of states, therefore, becomes

$$d(E) = \sum_n \delta (E - H(I)). \quad (21)$$

which can be recast as the sum of a smooth and an oscillatory term. The former originates from the Thomas-Fermi semi-classical approximation

$$d_{TF}(E) = \int \delta (E - \mathcal{H}(p, q)) \frac{dp dq}{(2\pi \hbar)^2}. \quad (22)$$

while the latter is obtained by expanding the effective action to quadratic order around the classical periodic orbits (Berry and Tabor, 1977; Emile et al., 2003):

$$d_{osc}(E) = \sum_{n=1}^{\infty} \frac{2\pi}{\hbar T_p} \left( \frac{S_n}{\hbar} - \mu_n + \frac{\pi}{4} \beta \right) \exp \left( i \pi \frac{\mu_n}{2} \frac{\beta}{4} \right), \quad (23)$$

where $Q_{ij} = \det (\mathcal{H}) \times \mathcal{H}_{ij}^{-1}$ is the co-matrix of $\mathcal{H}_{ij} = \partial_i \partial_j \mathcal{H}$, while $\beta$ is related to the signature of $\mathcal{H}_{ij}$.

We see, as in classical mechanics, one can also express the density of states as a sum of a smooth function $d_{TF}(E)$ and an oscillatory function which is defined on the periodic orbits of the semiclassical system. Due to the correspondence principle, we expect the Thomas-Fermi density of states to remain valid and only the oscillatory part to vary compared to the semi-classical derivation.

For non-integrable systems, however, the orbits no longer lie on invariant tori and a different method is needed for the evaluation of the trace

$$d(E) = -\frac{1}{\pi} \text{Tr} \left( \text{Re} (G_E(r, r)) \right) \quad (24)$$

where $G_E(r, r)$ is the Green-function associated with a given Hamiltonian $\mathcal{H}$. This new approach, based on the Green-function, was developed by Gutzwiller (1970, 1971). Here we shall not follow the details of the derivation, but only present the final, fully quantum mechanical expression for the density of states

$$d_{osc}(E) = \sum_{p.p.o.} \frac{T_p}{\pi \hbar} \sum_{n=1}^{\infty} \cos \left( \frac{n S_p - \frac{\pi}{2} \hbar p}{4} \right) \left| \det (M_p^2 - 1) \right|^{1/2} \quad (25)$$

where the summation runs over all primitive periodic orbits, and $M_p$ is the monodromy matrix for these primitive periodic orbits. Using this new method one can derive a semiclassical expression for the spectrum of a quantum system whose classical analogue is chaotic, when the usual Bohr-Sommerfeld quantisation rules cannot be applied. Gutzwiller’s result above, therefore, can be viewed as a bridge between the classical and quantum behaviour of a system, and can provide a rule as to how to quantise such a system. In this interpretation, Gutzwiller’s approach is similar to Feynman’s path integral description, where the quantum system is described in terms of an infinite sum over classical paths. For the interested reader we can suggest, without any reservation, Gutzwiller’s comprehensive book on classical and quantum chaos (Gutzwiller, 1991) and Brack and Bhaduri’s monograph giving an overview of semiclassical
physics (Brack and Bhaduri, 2003). In order to help the reader to visualise the emergence of periodic orbits in a quantum mechanical system we reproduce here a few quantum ‘scars’ from Heller’s numerical study. Figure 5 shows the probability distribution for three quantum eigenstates of the Bunimovich billiard. It is apparent how the isolated, unstable classical periodic orbits manifest themselves as paths along which the probability distribution is greatly enhanced. Gutzwiller’s idea to extract eigenvalues of a chaotic system via the periodic orbits, therefore, seems most plausible.

![Image of quantum scars](image_url)

**FIG. 5** Three eigenstates of the quantum stadium billiard are shown together with the major contributing unstable periodic orbits of the classical counterpart as thick solid lines. In the middle figure the guiding straight line for the \(\wedge\) shaped periodic orbit is omitted. From (Heller, 1984) with the kind permission of the author.

| Generic chaotic system | Riemann zeta function |
|------------------------|-----------------------|
| periodic orbit labels  | integers              | primes                |
| dimensionless action   | \(S_p/h\)             | \(T \ln(p)\)         |
| periods                | \(T_p\)               | \(\ln(p)\)           |
| stability factor\(^*\) | \(\det(M_p^n - 1)\)  | \(p^r\)              |
| Maslov index\(^*\)    | \(\mu_p\)            | \(2^r\)              |
| asymptotic limit       | \(h \to 0\)          | \(T_p \to \infty\)   |

\(^*\) Depending on how one maps the oscillatory part of the zeta zeros density (9b) onto Gutzwiller’s trace formula (25) the definition of the stability factor and the Maslov index can be different. Here we followed (Brack and Bhaduri, 2003), while another mapping can be found in (Berry and Keating, 1999b).

\(^\dagger\) Therefore the Maslov phase is \(\pi\), but this is not unique and one could also choose \(3\pi, 5\pi\), etc.

Below we pursue the proposed dynamics related to the Riemann zeta function.

In the early days of quantum mechanics Hilbert and Pólya suggested a physical way to verify Riemann’s Hypothesis:

> I spent two years in Göttingen ending around the begin of 1914. I tried to learn analytic number theory from Landau. He asked me one day: “You know some physics. Do you know a physical reason that the Riemann Hypothesis should be true.” This would be the case, I answered, if the nontrivial zeros of
The zeros of $\zeta(s)$ can be the spectrum of an operator, $\mathcal{R} = \frac{1}{2} \mathcal{I} + i \mathcal{H}$, where $\mathcal{H}$ is self-adjoint. This operator $\mathcal{H}$ might have an interpretation as a Hamiltonian of a physical system and, therefore, the key to the proof of the Riemann Hypothesis may have been coded in physics. Since the first occurrence of this conjecture a number of models have been promoted. Below we separate the models depending on whether they relate the zeros to the positive energy spectrum, i.e. the scattering states of a physical system, or to the negative energy spectrum, i.e. to the bound states of a quantum system.

1. Scattering state models

Let us first consider the possibility that the Riemann zeta function is associated with a quantum scattering problem.

A few decades after Riemann created a new geometry with his revolutionary work (Riemann, 1867), Hadamard examined the geodesics, the trajectories of freely moving bodies, on surfaces with negative curvature in detail (Hadamard, 1898) and noticed the occurrence of families of geodesics whose cross-section exhibits a fractal-like structure, as we would call it nowadays. These geodesics diverge exponentially, thus the distance between two trajectories, $\delta(t)$, however small initially, will grow exponentially, $\delta(t) \approx e^{\lambda t} \delta(0)$, where $\lambda$ is a positive number, called the Lyapunov exponent. This sensitivity of the system to the initial conditions, however, would not necessarily result in chaotic behaviour, provided the space for the trajectories is infinite. However, if the surface is compact, the trajectories cannot escape to infinity, rather mix on this surface. If one wishes to visualise a particular example, consider a donut with two holes. On this surface the trajectories remain bounded on the surface without the length of a geodesics being limited (Balazs and Voros, 1986; Bogomolny et al., 1995; Gutzwiller, 1991). These two properties, exponential sensitivity of the initial conditions and mixing, are the main requirements for chaotic motion (Cvitanović et al., 2014). The relative simplicity of the description of such surfaces with negative curvature, and the presence of completely chaotic classical motion motivated several authors in the mid-1980s to examine how such a system can be quantised, i.e. what properties do the solutions and eigenvalues of the equation $\mathcal{H}\phi = \lambda\phi$ possess.

More precisely, for free motion, one seeks the solution of

$$-\Delta \phi_n = \lambda_n \phi_n$$

where $\phi_n$ are required to be square integrable and the appropriate boundary conditions are also provided. Over a compact domain equation (26) has only discrete eigenvalues. On a surface with negative curvature, the non-euclidean Green’s theorem shows that the eigenvalues must have the form $\lambda_n = \frac{s}{2} + i\rho_n$ ($s$ is real) (Gel’fand and Pjatezkii-Shapiro, 1954). This resemblance immediately suggests a connection with the zeros of the Riemann $\zeta(s)$. It is also proven that, for a compact surface, the set of $n$’s is finite, but for a non-compact surface, a continuous part of the spectrum can also appear. In the latter case the scattering (continuous spectrum) is non-conventional, because it is the result of the geometry (curvature, compactedness) and not the physical interaction between particles.

In order to express the eigenvalue density the Green-function is needed. Interestingly, on a surface with negative curvature the Green-function can be explicitly written as a sum of individual Green-functions corresponding to the periodic orbits. It is also a fact that, all periodic orbits are unstable and their action is $S(E) = k\ell$, where $k$ is the momentum related to the energy by $2mE/k^2 = k^2 + 1/4$ and $\ell$ defines the length of a closed geodesic belonging to a given conjugacy class. In this geometry the density of states is expressed by the Selberg trace formula (Selberg, 1949)

$$\bar{\rho}(k) = \frac{A}{2\pi} k \tanh (k\pi) + \frac{1}{2\pi} \sum_{|p|} \sum_{n=1}^{\infty} \frac{\ell_p \cos(nk\ell_p)}{\sinh(n\ell_p/2)}$$

where $A$ is the area of the surface, the first summation runs over conjugacy classes of primitive elements $p$, the second, their repetitions. It is important to note, the Selberg trace formula holds exactly, in contrast to other trace formulae, because no semi-classical approximation has been applied, although its convergence property is similar to the Gutzwiller form: for large $k$ the Selberg and Gutzwiller trace formulae converge, since the metric is locally Euclidean and waves with short wavelength lose their sensitivity to the local curvature of the metric. In this system, the transient scattering states were examined by Pavlov and Faddeev who related the nontrivial zeros of the zeta function to the complex poles of the scattering matrix (Pavlov and Faddeev, 1975):

$$S(k) = \pi^{-2ik} \frac{\Gamma(\frac{1}{4} + ik)\zeta(1 + 2ik)}{\Gamma(\frac{1}{4} - ik)\zeta(1 - 2ik)}$$

6 G. Pólya refers here to the Riemann $\zeta(s)$ function.
7 See the scanned pages on Odlyzko’s personal website: http://www.dtc.umn.edu/~odlyzko/polya/
Despite this natural occurrence of the Riemann zeta function and its non-trivial zeros, no further insight into the zeros has been gained via this route. Detailed discussion of the Selberg trace formula can be found in [Hejhal, 1976, 1983] or more physics oriented approaches in [Stöckmann, 1999] and [Wardlaw and Jaworski, 1989] and in the context of the Casimir-effect in [Elizalde, 1993; Kurokawa and Wakayama, 2002; Schaden, 2006].

So, let us return to the scattering formalism in the standard Euclidean space. Joffily, motivated by Pavlov and Fadeev (Pavlov and Fadeev, 1975), examined the scattering states of a non-relativistic, spinless particle under the influence of a spherically symmetric, local and finite potential. He examined the Jost solutions of this scattering problem (Joffily, 2003), which differ from the physical solution of the Schrödinger equation in their asymptotics (Alfaro and Regge, 1965; Newton, 1982). In standard non-relativistic scattering theory the S-matrix is given by

$$ S(k) = e^{2\delta(k)} = \frac{f_+(k)}{f_-(k)} $$

where $\delta(k)$ is the phase shift, and $f_\pm(k)$ are the Jost solutions defined by their boundary conditions $\lim_{r \to \infty} (f_\pm(k)e^{\mp ikr}) = 1$ (Alfaro and Regge, 1965). Provided the potential has a finite range and decreases sufficiently rapidly, the Jost solution $f_+(k)$ is proven to have infinitely many zeros, corresponding to the solutions of the Schrödinger equation as outgoing or incoming waves. Resonances (i.e. states with finite lifetime) occur if $S(k)$ has poles on the complex $k$ plane with negative imaginary parts: $k_n^2 = \epsilon_n - i\Gamma_n / 2$, where $\epsilon_n$ and $\Gamma_n$ stand for the energy and inverse lifetime associated with the $n$th state. Joffily introduces a mapping between these zeros of $f_+(k)$ onto the critical line and shows they coincide with the non-trivial zeros of the Riemann zeta function. He associates this artificial system with a vacuum and the zeros are interpreted as an infinity of virtual resonances, and thus reflect the chaotic nature of the vacuum (Joffily, 2003, 2004). This interpretation has also been extended using relativistic scattering (Joffily, 2007).

In another scattering based approach, Chadan and Musette analysed the so-called ‘coupling constant spectrum’ of a radially symmetric three-dimensional Hamiltonian (Chadan and Musette, 1993) where the potential is chosen from a singular family of functions

$$ \mathcal{H}_{CM} = -\frac{d^2}{dr^2} - \ell(\ell+1) / r^2 + \frac{1}{r^2}f_{CM} $$

where $f_{CM}$ has logarithmic singularities at $r = 0$. They argued that the coupling constant spectrum coincides ‘approximately’ with the non-trivial Riemann zeros if the problem is restricted to a finite, closed interval $r \in [0, e^{-4\pi/\ell}]$. Mathematically rigorous detailed analysis and the extension of the potential family was carried out by (Khalf, 2002). Furthermore, the existence of a three dimensional potential, $U_P(r)$, was derived whose $s$-wave scattering amplitude has the complex zeros of the Riemann zeta function as “redundant poles”. Examination of $\zeta(s)$ using a quantum scattering approach is further motivated if one compares the plot of the phase of $\zeta(s)$ on the complex plane with with the usual Argand-diagram of the scattering amplitude corresponding to a collision.

As a specific example, for completely elastic collisions, the scattering amplitude should be a perfect circle on the complex plane with unit radius centred on $(0,1)$. For inelastic collisions this circle deforms. The phase of $\zeta(s)$, after interchanging the roles of the real and imaginary axes, qualitatively resembles the Argand diagram of a scattering amplitude. This geometric similarity suggests an analysis of $\zeta(s)$ as if it represented the scattering amplitude of a real collision of particles. This analogy, however, is not perfect since $\zeta(s)$ does become negative while

---

8 The Jost functions are the solutions of the Schrödinger equation with the following asymptotic behaviour:

$$ \lim_{x \to \infty} (e^{ikx} f(\lambda, k, x)) = 1 $$

where $\lambda = \ell + \frac{1}{2}$ is the shifted angular momentum, $k \sim \sqrt{E}$ and $x \in (-\infty, \infty)$. This choice of the boundary condition is motivated by our physical picture, i.e. the particle should be represented by free plane waves far from the local potential. The real physical solution of the Schrödinger equation can be expressed as a linear combination of the two Jost functions.

9 Argand diagrams can be thought of as a parametric plot of the inherently complex scattering amplitude on the complex plane, and the collision energy plays the role of the parameter (see for example (Bohm and Loewe, 2001)) or (Bhaduri, 1988).
the Argand diagram of the scattering amplitude corresponding to a realistic collision does not. Bhaduri advocates neglecting these small differences which do not affect their most important result, namely the phase $\theta(t)$ of the Riemann zeta function along the critical line, $\zeta(1/2 + it) = Z(t)e^{-\gamma t}$, is intimately connected to the quantum scattering of a particle on a saddle-like surface (Bhaduri et al., 1995, 1997).

To illustrate this, let us consider a non-relativistic particle moving in an inverted harmonic oscillator potential along the half-line ($x \geq 0$). The Schrödinger equation reads as

$$-rac{\hbar^2}{2m} \frac{d^2}{dx^2} \Phi(x) - \frac{1}{2} m \omega^2 x^2 \Phi(x) = E \Phi(x) \quad (31)$$

where we require that $\Phi(x = 0) = 0$. This problem can be mapped onto a repulsive Coulomb problem of which the phase shift $\delta(t)$ can be exactly expressed (Flügge, 1974). The oscillatory part of the phase shift is given by

$$\delta(t) = \delta_{\text{smooth}}(t) +$$

$$\Im \left[ \ln \left( \frac{1}{4} + \frac{it}{2} \right) \right] - \ln \left( \frac{1}{4} - \frac{it}{2} \right) \right] \quad (32)$$

which is exactly the phase of the Riemann zeta function. Two years after their first result, Bhaduri et al. extended this one-dimensional model to a two-dimensional one where in one direction the potential is a traditional confining parabolic potential, and in the perpendicular direction ($y$) they kept the inverted harmonic oscillator (Bhaduri et al., 1974). This choice was motivated by the analysis of the Gutzwiller trace formula on the $\sigma = 1$ border of the critical line, and also by the form of the electrostatic potential at the bottleneck of a quantum contact in a mesoscopic structure (Büttiker, 1990).

While the inverted oscillator reproduced the oscillating part of the $\zeta(s)$ phase in Bhaduri’s work, Berry and Keating showed that a regularisation of a surprisingly simple one-dimensional classical Hamiltonian, $\mathcal{H} = xp$, reproduces the smooth counting function of the zeros (Berry and Keating, 1999a). We note here that this choice of $\mathcal{H}$ is a canonically rotated form of the inverted oscillator Hamiltonian $\sim (p^2 - x^2)$. Moreover, the quantum mechanical model of the corresponding symmetrised Hamiltonian, $\mathcal{H} = (xp + px)/2$ has also been investigated and exactly solved preserving the self-adjoint property of the Hamiltonian (Sierra, 2007; Twamley and Milburn, 2000). The beauty of the $xp$-or inverted oscillator model is that it satisfies most of the properties listed earlier (see page 9): valid as a classical mechanical model; the dynamics is one-dimensional and uniformly unstable since the solution of the Hamiltonian equations are exponentially decaying or diverging; it lacks time-reversal symmetry. However, the trajectories are not bounded causing significant hardship in the semiclassical quantisation. As in the hyperbolic case, the boundary conditions or the way the phase space is regularised/ compactified become decisive. Berry and Keating suggested a simple regularisation (Berry and Keating, 1999a) by introducing a cut-off in both position and momentum. This process results in a finite area, which can be filled up with Planck-cells of size $\hbar$, thus counting the number of available quantum states. Another approach is available if one notices the dilation symmetry of the Hamiltonian ($x \mapsto \lambda x$, and $p \mapsto p/\lambda$). This symmetry manifests itself in the transformation of the wavefunction as

$$\psi(x) = \frac{1}{\lambda^{1/2 - iE}} \psi(x) \quad (33)$$

and one might suggest restricting ourselves to $\lambda$ being a positive integer. This could be an attractive suggestion, because the wave-packet, generated by the uniform superpositions of all these transformed wavefunctions is

$$\Psi(x) = \sum_{\lambda=1}^{\infty} \psi(\lambda x) = \zeta \left( \frac{1}{2} - iE \right) \psi(x). \quad (34)$$

However, there is no physical motivation which would require this $\zeta$ pre-factor to vanish. Furthermore this integer-based dilation-symmetry does not form a group, because the multiplicative inverse element (which would be $\lambda = 1/m$) is missing.

Berry and Keating also established a peculiar canonical transformation ($X = 2\pi/p, P = xp^2/2\pi$) for this Hamiltonian, which exchanges and mixes the roles of the physical position and momentum, but was uncertain “how to convert this “quantum exchange” into an effective boundary condition” (Berry and Keating, 1999). Aneva also analyses this boundary condition for a hyperbolic dynamical system with conformal geometry and shows how this exchange transformation arises as a result of boundary conditions (Aneva, 1999, 2001a,b).

Later Sierra generalised Berry’s model in two different ways, first by incorporating the fluctuation terms, $d_{\text{osc}}(E)$ (see relation (9b)), via changed boundary conditions (Sierra, 2008). Together with Townsend, they also considered the motion of a charged particle (electron) moving on the $[x,y]$ plane in a constant uniform perpendicular magnetic field, and in an electric potential described by the following Hamiltonian

$$\mathcal{H} = \frac{1}{2\mu} \left[ p_x^2 + \left( p_y + \frac{eB}{c} x \right)^2 \right] + e\lambda xy. \quad (35)$$

In this model, the number of semiclassical quantum states with energy less than $E$ has the same functional form as the counting function of the $\zeta(s)$ zeros, eq. [10] i.e. the smooth part of the Riemann-zeros is reconstructed by the lowest-lying Landau level of the charged particle. The fluctuation term – as they speculate – might be explained by the contribution of higher Landau-levels. This surprise, however, is only supported by estimating the order of magnitude of these higher contribution
and comparing it to that of the Riemann \( \zeta(s) \). This model has the additional attraction of being potentially accessible to experimentalists, including in lower spatial dimensions (Li and Andrej 2007, Park et al. 2009, Toet et al. 1991).

Exploiting the \( x \leftrightarrow p \) exchange symmetry of this model and using the Riemann-Siegel formula for the \( \zeta(s) \) function, Sierra created a new model in which the Jost solutions are directly proportional to the Riemann zeta function, and the non-trivial zeros become the energies of the bound states. This achievement does not, however, prove the Riemann Hypothesis, as Sierra explicitly states “we cannot exclude the existence of zeros outside the critical line”.

In summary, we first introduced, motivated by Gutzwiller’s trace formula, a quantum mechanical model on a surface with negative curvature, which lead us to the mathematically exact Selberg trace formula. The importance of this result is at least twofold. Firstly, it reassures us that describing chaotic systems via the periodic orbits is likely to be feasible, and secondly demonstrates the role of periodic orbits in a generic system in determining the smooth and fluctuating parts of the density of states. We further elaborated on another non-Euclidean model, proposed by Pavlov and Fadeev, in which the Riemann-zeta function determines the S-matrix over the complex energy plane.

Converting these results into the usual Euclidean space, however, seems challenging. Although a few models have successfully reproduced the smooth part of the density of quantum states, the fluctuation terms of these models differ from that of the Riemann zeta function.

2. Bound state models

From the 1950s a new approach, the Random Matrix Theory, emerged from the study of the spectrum of heavy nuclei. The same statistical apparatus had also been used to analyse the statistical properties of the seemingly random Riemann zeta zeros, and lead to the conjecture that the \( \zeta(s) \) zeros belong to one particular universality class (Bohigas et al. 1984a, 1986), the so-called Gaussian Unitary Ensemble (see later in section III.C). This result suggested property 3 on Berry’s list. However, Wu and Sprung generated a one-dimensional, therefore integrable, quantum mechanical model which can possess the Riemann zeta zeros as energy eigenvalues (Wu and Sprung 1993) and show the same level-repulsion as that observed in quantum chaos. This was a contradictory result since on one hand the \( \zeta(s) \) zeros follow a statistics specific for systems violating time-reversal symmetry, on the other hand, Wu and Sprung’s model, by definition, was invariant under time reversal. However, the proposed model was not lacking in irregularity, since the potential reproducing the Riemann zeta zeros appeared to be a fractal, a self-similar mathematical object. Nevertheless, these authors derived for the first time a smooth, semi-classical potential which generate the smooth part of \( N(E) \)

\[
N(E) = \frac{1}{h} \iint_{\mathcal{H} \leq E} dx \, dp = \frac{2}{\pi} \int_{0}^{c_{\text{max}}} \sqrt{E - V(x)} \, dx \tag{36}
\]

with the \( 2m/h^2 \) set to unity. Solving this Abel-type integral equation one may derive the following implicit expression for \( V(x) \)

\[
x(V) = \frac{1}{\pi} \left[ \sqrt{V - V_0 \ln \left( \frac{V_0}{2\pi} \right)} + \sqrt{V \ln \left( \frac{\sqrt{V + \sqrt{V - V_0}}}{\sqrt{V - \sqrt{V - V_0}} - V_0} \right)} \right] \tag{37}
\]

where \( V_0 \) has to be chosen such that the potential is not multi-valued, i.e. \( V_0 \leq 2\pi \). The choice of \( V_0 \) affects the potential at its bottom (\( x \approx 0 \)), but for large \( x \) it does not have a significant impact and for \( x \gg 1 \)

\[
x(V) = \frac{\sqrt{V}}{\pi} \ln \left( \frac{2V}{\pi \epsilon^2} \right) \tag{38}
\]

(see figure 1 in Wu and Sprung 1993). We note here that Mussardo, using similar semiclassical arguments as Wu and Sprung, recently also gave a simple expression for a smooth potential supporting the prime numbers (Mussardo 1997) as energy eigenvalues. Furthermore, Mussardo also proposed a hypothetical resonance experiment to carry out primality testing: this theoretically infinite potential could be truncated at some high energy; thus transforming it into a finite well. If an incident wave radiated onto this well has energy \( E = n\hbar \omega \) where \( n \) is a prime number, then it should cause a sharp resonance peak in the transmission spectrum - argues Mussardo.

Turning back to the smooth potential studied by Wu and Sprung, which is able to ‘roughly’ reproduce \( N(E) \), it is then modified to have the low lying \( \zeta(s) \) zeros exactly. In order to achieve this goal Wu and Sprung set up a least-square minimisation routine, to minimise the difference between the actual energy eigenvalues and the exact zeros. The result was surprising, since the potential curve became coarse and resembled a random potential. They analysed this curve using the standard box-counting technique and measured a \( D = 1.3 \) fractal dimension for the potential reconstructing the Riemann zeta zeros.

Ramani et al. pointed out (Ramani et al. 1995) that the apparent contradiction between Berry’s conjecture and Wu and Sprung’s model, i.e. whether or not the physical system exhibits time-reversal symmetry, is caused by the coarse curve of the potential, since any

\[\text{\textsuperscript{10}}\text{In the mathematical literature the argument is usually denoted by } T, \text{ as in section } \text{III.}}\text{Motivated by physics, we use here } E.\]
smooth one-dimensional potential would lead to locally evenly spread energy levels, which is not the case for Wu’s potential. They also provided a very efficient algorithm, the “dressing transformation” with which one can build up the quantum potential from individual energy eigenvalues. However, they standardised the spectrum using the ‘spectrum unfolding’ technique which eventually lead them to the conclusion: the fractal dimension of the potential supporting the Riemann zeta zeros has \( d \to 2 \) rather than that measured by Wu and Sprung. In a reply (Wu and Sprung, 1995), Wu and Sprung pointed out that this difference in fractal dimension is putatively caused by the alternative choice of spectrum. As they argued, Ramani’s spectrum does not have the same average density, long range correlation and nearest-level spacing distribution as the Riemann zeta function, therefore one cannot draw valuable conclusions regarding the potential.

Nearly a decade after Wu and Sprung’s original article, van Zyl and Hutchinson attempted to clarify the questions raised by the two previous works (van Zyl and Hutchinson, 2003). They showed that for the same set of energy levels different potential generating techniques (the variational approach used by Wu and Sprung, the dressing-transformation used by Ramani et al.) lead to the same potential, depicted in Figure 7. This result had been further strengthened by Schumayer et al. who used the inverse scattering transformation as a third technique obtaining the same potentials as in the earlier works (Schumayer et al., 2008). It is noteworthy to mention that the inverse scattering transform guarantees the uniqueness of the potential in one-dimension. This analysis, therefore, elucidated that the difference in measured fractal dimension cannot originate from the method of inversion. Moreover, they confirmed \( d = 1.5 \) for the Riemann zeta potential. These works all demonstrated the importance of long-range correlations in determining the fractal dimension of the potential.

In a similar manner to that of the the Riemann \( \zeta(s) \) zeros, the prime numbers can also be considered as an energy spectrum, thus a potential can be associated with them and it also proves to be fractal, but with a larger fractal dimension, \( d = 1.8 \). This result is somewhat puzzling. The two sets, those of the zeta zeros and the prime numbers can be mapped onto each other via eq. 10, but the nearest-neighbour spacing distribution of prime numbers is known to be Poisson-like (almost uncorrelated random distribution) while that of the Riemann zeros is rooted in the Gaussian Unitary Ensemble, and exhibits the corresponding correlations (see expression 10 in section III.C). One may, therefore, conclude that Riemann’s formulae converts two very different random distributions into each other, or as Sakhr et al. put it (Sakhr et al., 2003): “it is possible to generate the almost uncorrelated sequence of the primes from the interference of the highly-correlated Riemann zeros”.

Regarding the fractal nature, Schumayer et al. also established that the potentials associated with either the zeros of \( \zeta(s) \) or with the prime numbers are multifractals, i.e. these potential curves cannot be characterised by one number \( d \), but a range of dimension is necessary to describe their properties (for definition see Schumayer et al., 2008).

Finally, at the end of this section devoted to the quantum mechanical models of the Riemann zeta function, we briefly refer to another alternative spectral interpretation of the zeros proposed by Connes (Connes, 1999). During the comparison of the Gutzwiller’s trace formula for quantum mechanical systems and that of the \( \zeta(s) \) function we noticed the overall sign difference in \( d_{osc} \) (see the negative sign in equation 10 in front of the summation), i.e. the contribution of the periodic orbits should be subtracted and not added to the smooth density of states, \( d(T) \) (Berry, 1986). This sign difference led Connes to interpret the zeros as gaps, missing lines from the otherwise continuous energy spectrum rather than discrete energy levels.

C. Nuclear physics

Random Matrix Theory (RMT) has been successfully applied to predict ensemble averages of observables for heavy nuclei. Even though the Riemann zeros are distributed randomly, some of their statistical quantities correspond to that of the Gauss Unitary Ensemble. We discuss the RMT briefly for historical reasons. The reason for brevity owes to two recent Colloquia devoted to RMT (Papenbrock and Weidenmüller, 2007; Weidenmüller and Mitchell, 2008).

Unfortunately the degrees of freedom of even a moderately large nucleus are still far beyond our computa-
tional capability, be it analytical or numerical. Similar problems, although the number of components are on a different scale, have occurred before in physics and engendered the development of a new branch of physics, statistical mechanics. This is exactly what Wigner had in mind when he suggested a statistical description of nuclei (Wigner, 1951). He suggested that nuclei can be statistically described using random matrices carefully chosen from pre-determined ensembles. The new description emerging from this examination is the Random Matrix Theory.

Although random matrix theory emerged from the statistical description of nuclei, it has already infiltrated into many different areas of physics. Recent developments of this branch of physics have been reviewed in (Bohigas, 1989; Forrester et al., 2003; Weidenmüller and Mitchell, 2009). Moreover we can suggest the monograph by one of the leading figures of random matrix theory (Mehta, 2004).

But how to choose the ensemble of random matrices suitable for a certain system, or for the Riemann ζ(s) function? Throughout classical mechanics symmetry plays a decisive role in determining the dynamics of different systems. If a physical system has a symmetry it implies, via Noether’s theorem, the existence of a conserved quantity, e.g. the translational invariance in time dictates energy conservation, continuous rotational invariance requires the angular momentum remain constant. These symmetries limit the possible forms of the Hamiltonian describing the given system. Therefore, if one wants to approximate a Hamiltonian with a large, but finite dimensional matrix these symmetries will determine the type and structure of the matrix, whether it is real or complex, symmetric or hermitian (Dyson, 1962).

In the case of an integrable system, the conserved quantities are all known. Therefore the Hamiltonian can be diagonalised, with each eigenvalue forming its own symmetry-class. This leads to the assumption that these eigenvalues are completely uncorrelated. Let us also assume that the average spacing between eigenvalues is unity in the overall sequence of eigenvalues. If p(s) denotes the probability distribution of nearest neighbour spacings, i.e. if ε₁ and ε₂ are eigenvalues of the given system, then ε₁ − ε₂ = s, then one can express (Stöckmann, 1999) the probability of finding two eigenvalues in a distance between s and s + ds with no other eigenvalues in between. Dividing the distance s into N equal intervals, the probability is simply

\[ p(s) ds = \lim_{N \to \infty} \left( 1 - \frac{s}{N} \right)^N ds \quad \text{(39)} \]

In the \( N \to \infty \) limit the right hand side becomes the exponential function. Therefore the probability distribution \( p(s) = \exp(-s) \) is the Poisson distribution with parameter equal to 1. This is quite a general result for integrable systems as Berry and Tabor have demonstrated (Berry and Tabor, 1977). Similarly, one can deduce similar probability distributions for universality classes of random matrices, e.g. Gaussian Unitary Ensemble, Gaussian Orthogonal Ensemble, etc. The classification refers to the universality conjecture: if the classical dynamics is integrable then \( p(s) \) corresponds to the Poisson ensemble, while in the chaotic case \( p(s) \) coincides with the corresponding quantity for the eigenvalues of a suitable ensemble of random matrices (Bohigas et al., 1984a). Furthermore, the local statistics of the eigenvalues converge as the order of the matrix increases.

How is this connected to the Riemann-zeta zeros? The zeros can be treated as eigenvalues of a fictitious physical system, just as Hilbert and Pólya suggested, and their statistical properties examined. In 1973 Hugh Montgomery showed (Montgomery, 1973) that the pair-correlation function of the zeros is

\[ r_2(x) = 1 - \frac{\sin^2(\pi x)}{\pi x} \quad \text{(40)} \]

provided the Riemann Hypothesis is true. Freeman J. Dyson, during an informal discussion over tea (Cipra, 1999), pointed out to Montgomery that this is exactly the same result as obtained for random matrices picked from the Gaussian Unitary Ensemble. However, this statement is made in the asymptotic limit, i.e. as one goes to infinity on the critical line, \( 1/2 + iE \), or in the RMT language as the size of the matrices, \( N \), tends to infinity. At finite height \( E \) or dimensionality \( N \) discrepancies may occur compared to expression (40). Interestingly, it was shown using heuristic arguments, that the nearest-neighbour spacing distribution of the zeta zeros and that of unitary random matrices of finite dimension are the same (Bogomolny and Keating, 1995). Moreover, the same authors extended their study of correlation functions (Bogomolny and Keating, 1996), \( r_n \), of order \( n \) (\( n \geq 2 \)) and proved, in the appropriate asymptotic limit, \( r_n \) of the Riemann zeta zeros are equivalent to the corresponding GUE result. This result was complimentary to Montgomery’s second order (Montgomery, 1973), Hejhal’s third order (Hejhal, 1994) and Rudnick and Sarnak’s general result for the \( n \)th order correlation function (Rudnick and Sarnak, 1999).

What does this result demand from a model of the Riemann zeros? The striking similarity between the pair-correlation function of the \( \zeta(s) \) zeros and the eigenvalues of random matrices from the GUE ensemble only holds for short-range statistics. Odlyzko, by calculating the statistics for substantial numbers of zeros, carried out an empirical test (Odlyzko, 1987) and confirmed Berry’s predictions (Berry, 1985) about the discrepancies between the GUE theory and computed behaviour of the \( \zeta(s) \) zeros. The long-range correlation and the small spacing statistics of the \( \zeta(s) \) zeros noticeably deviate from the GUE prediction. This is expected (Berry, 1989a).
since long-range correlations are dominated by the short periodic orbits, which are system specific and therefore not universal. For $\zeta(s)$ the mean separation between zeros is $\ln(E/2\pi)$ while the smallest period is $\sim \ln(2)$ (Table I). Conclusively, the GUE-predicted universal correlation for zeros near $E$ should fall beyond $\ln(E/2\pi)/\ln(2)$ (Berry and Keating, 1999a). Despite the deviation explained above, the statistics of the $\zeta(s)$ zeros asymptotically coincide with those of the GUE ensemble, consequently the corresponding quantum system ought to violate time-reversal symmetry (Berry and Keating, 1999a,b). This may have motivated Berry and Keating’s choice of a $\sim (xp + px)$ as a Hamiltonian.

Finally, we must mention an unexpected spin-off result of random matrix theory related to the Pólya-Hilbert conjecture. Crehan asserted (Crehan, 1995) that for any bounded sequence there are infinitely many classically integrable Hamiltonians for which the corresponding quantum spectrum coincides with this sequence. Furthermore, as an example for his theorem, he shows that infinitely many classically integrable non-linear oscillators are capable of exactly reproducing the Riemann-zeta zeros when they are quantised. Unfortunately, the theorem is an existence theorem and not a constructive one. If such a system could be created, whether physically or just theoretically, that would be aesthetically pleasing: it would connect the most studied physical model (oscillator) with the basis of our arithmetic (prime numbers). Crehan’s result is promising and is also supported by the relationship between the Riemann $\zeta(s)$ zeros and the Painlevé V equation, the latter of which plays a central role in the theory of completely integrable dynamical systems (Ablowitz and Clarkson, 1991).

Finally, in this section we briefly mention the notion of quantum ergodicity which attracted substantial attention in the last three decades in the search for links between classical and quantum ergodicity, i.e. what “fingerprint” the classical chaos leaves in the physical properties if we quantise the system, especially in the long-time behaviour. Only few rigorous results (Schreiman, 1974; de Verdiére, 1985; Zelditch and Zworski, 1996) are known, and one of them says that the expectation value of operators over individual eigenstates is almost always the ergodic, microcanonical average of the classical version of the operator. However, the theoretically rigorous understanding of quantum ergodicity is still in its infancy. Numerical simulations suggest though that quantum chaotic systems exhibit universal behaviour at a particular length scale, and at this scale the statistics of the eigenvalues resemble that of large random matrices chosen from specific ensembles (Agam et al., 1995; Berry, 1977; Bohigas et al., 1984b; Gutzwiller, 1991; Heller, 1984). It is unfortunate that this length scale is so minute that it hinders the numerical simulations substantially. Nevertheless, it has also been shown theoretically (Kaplan and Heller, 1996; Tomsovic and Heller, 1991) that quantum eigenstates must deviate from the RMT predictions. These corrections may stand out from the spread out background of RMT, just as the unstable periodic orbits do as eigenstates with enhanced amplitudes as depicted in Figure 5.

Although further numerical simulations (Bäcker et al., 1998; Kaplan and Heller, 1999) provide some evidence regarding the connection between RMT and quantum ergodicity, its interpretation and strength remain open questions.

### D. Condensed matter physics

In condensed matter physics the fundamental structure is the crystal lattice. Below we examine the connection of the lattice with the generalised Riemann hypothesis. We also show how the specific heat capacity of a solid restricts the location of the $\zeta(s)$ zeros.

One of the fundamental bases of modern condensed matter physics is the geometrical structure of solids; the lattice. The examination of this mathematical structure is necessary to understand even the basic properties of matter. The regular structure of a perfect lattice is suitable for immediate comparison with regularities among the natural numbers, and therefore it is not a surprise that many number-theoretical functions arise in crystallography, e.g. Ninham et al. present a witty review on the Möbius function (Ninham et al., 1992). For those mathematically more inclined we suggest the book “From Number Theory to Physics” by Waldschmidt (Waldschmidt et al., 1995). Moreover, not only the perfect regularity of a lattice, but also the lack of this regularity can be related to the Riemann zeta function, as Dyson indicated recently (Dyson, 2009): “A fourth joke of nature is a similarity in behavior between quasi-crystals and the zeros of the Riemann Zeta function.” In the following, we briefly examine why a solid state constituted by ions should even exist, what binds these ions to each other?

Ions arrange themselves into a structure which maximises the attractive interaction between unlike and minimises the repulsive interaction between like charges. In an ionic crystal, such as NaCl, the main contribution to the binding energy has an electrostatic origin with the van der Waals term only a few percent of the former. The electrostatic term is called the Madelung energy, and the energy of one ion in the solid is called the Madelung constant.

For the sake of simplicity, let us first imagine a one-dimensional infinitely long ionic lattice. Cations and anions are located next to each other at a distance $a$, in a simplified NaCl structure. If simply two unit charges $q$ were positioned at the same distance $a$, the electric potential energy of one of the charges would be $U = q^2/4\pi\varepsilon_0 a$. In a solid each ion is in the field of all the remaining charges, both positive and negative. The
total electrostatic potential energy of one ion at position \(i\) in the lattice is therefore

\[
U_i = \sum_{j \neq i} \frac{1}{4\pi\varepsilon_0} \frac{(-1)^{i-j} q^2}{|i-j|} = \frac{1}{4\pi\varepsilon_0} q^2 \sum_{k \neq 0} \frac{(-1)^k}{k}.
\]  

(41)

where \(j\) runs over all lattice sites except \(i\) in the first summation, and in the second we have changed the running variable to \(k = |i-j|\). In a finite lattice we have 2\(N\) ions, but in (41) each term belongs to two ions, therefore the total electrostatic potential energy of the finite lattice is

\[
U_{\text{total}} = \frac{1}{2} 2N U_i = N \frac{1}{4\pi\varepsilon_0} q^2 \sum_{k \neq 0} \frac{(-1)^k}{k}.
\]  

(42)

This form of \(U_{\text{total}}\) can be divided into three terms: \(N\) which guarantees the extensive nature of the energy, an energy factor, \(q^2/4\pi\varepsilon_0 a\), and also a numerical factor depending only on the lattice structure. One sees directly that the influence of the lattice on the total electrostatic energy is comprised of an infinite sum. Since this energy term has to be negative in order to describe binding, we incorporate this sign into the Madelung constant \(\alpha_{1D}\) as

\[
\alpha_{1D} = 2 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}.
\]  

(43)

where the factor 2 appears because of the mirror-symmetry around the \(i\)th ion. The total energy can be written as \(U_{\text{total}} = -\alpha_{1D} N q^2/4\pi\varepsilon_0 a\), which is negative if \(\alpha_{1D} > 0\).

Generalising the NaCl structure we examined above for the realistic three-dimensional case, one can write the Madelung constant for this lattice as

\[
\alpha_{3D} = \sum_{(i,k,l) \neq (0,0,0)} \frac{(-1)^{i+j+k+1}}{(i^2 + j^2 + k^2)^{1/2}}.
\]  

(44)

Although it is tempting to evaluate this summation by approximating the terms on concentric spheres centred at the reference ion \((i=j=k=0)\) and utilising the symmetry, the resulting series, \(6-12/\sqrt{2}+8/\sqrt{3}-\ldots\) is divergent which is physically unsatisfactory. The convergence properties of such sums have been extensively investigated (Borwein et al., 1985; Chaba and Pathria, 1975; 1976a,b, 1977). The sum (44) is an alternating and conditionally convergent sum. The denominator of the summand is a quadratic form, therefore the Madelung constant for a simple cubic structure can be formally written as \(\zeta_{EP}(1/2,\delta_{m,n})\) where \(\zeta_{EP}\) is the Epstein zeta function (see below), and \(m, n = 1, 2, 3\). The second argument, \(\delta_{m,n}\), is determined by the type of the lattice, and in crystallography it is a quadratic form \(P\) given by the Gram matrix \(p_{mn} = e_m e_n\), where \(e_m\) is the \(m\)th lattice vector. Therefore, for example, the Madelung constant for the body-centered cubic structure can be formally written as

\[
\alpha_{bcc} = \zeta_{EP}\left(\frac{1}{2}, \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\right) = 1.762675.
\]  

(45)

Here we have only dealt with the pure Coulomb-interaction, but this treatment can be extended to screened electrostatic interactions as well (Kanemitsu et al., 2003).

The infinite sum in (43) strongly resembles the Riemann-zeta function, except each term is weighted by a factor \((-1)^{k+1}\), and its numerical value is \(\alpha_{1D} = 2 \ln(2) \approx 1.3863\). Although in two and three dimensions the summation can be written explicitly, obtaining a precise numerical value is far from easy and the Epstein-zeta function is required. This function can be thought of as a generalised zeta-function (Ivic, 2003; Shanker, 2006) which is defined by

\[
\zeta_{EP}(s, P) = \sum_{P \neq 0} \frac{1}{P^s}.
\]  

(46)

where \(P\) is a quadratic form defined on a \(d\)-dimensional lattice. All lattice points for which \(P \equiv 0\) are excluded from the summation. This function can be analytically continued to the same domain as the Riemann-zeta function and also has its only pole at \(s = 1\) with residue \(\pi/\sqrt{\Delta}\). The similarity goes further since \(\zeta_{EP}(s, P)\) also satisfies a functional equation expressing mirror symmetry. Thus, there is an inclination to generalise Riemann’s conjecture: all non-trivial zeros of \(\zeta_{EP}(s, P)\) have real part one half. The temptation to do so is strengthened if one chooses specific quadratic forms, e.g. \(\zeta_{EP}(s, \text{Id}_1) = 2\zeta(2s)\), or \(\zeta_{EP}(s, \text{Id}_4) \sim \zeta(s)\zeta(s - 1)\), where \(\text{Id}_n\) is the \(n\)-dimensional identity matrix. Indeed, it was shown eighty years ago that for binary quadratic forms (two dimensional lattice), infinitely many zeros of \(\zeta_{EP}(s, P)\) lie on the critical line (Potter and Titchmarsh, 1935) in a similar manner to Hardy for the Riemann zeta function (Hardy, 1914). Remarkably, however, it has also been shown that in any dimension one can construct such a \(P\), that the generalised hypothesis does not hold (Terras, 1980).

This, admittedly negative, result shows the intriguing properties of such sums have been extensively investigated (Borwein et al., 1985; Chaba and Pathria, 1975; 1976a,b, 1977). The sum (44) is an alternating and conditionally convergent sum. The denominator of the summand is a quadratic form, therefore the Madelung constant for a simple cubic structure can be formally written as

\[
\zeta_{EP}(1/2, \delta_{m,n})\] where \(\zeta_{EP}\) is the Epstein zeta function (see below), and \(m, n = 1, 2, 3\). The second argument, \(\delta_{m,n}\), is determined by the type of the lattice, and in crystallography it is a quadratic form \(P\) given by the Gram matrix \(p_{mn} = e_m e_n\), where \(e_m\) is the \(m\)th lattice vector. Therefore, for example, the Madelung constant for the body-centered cubic structure can be formally written as

\[
\alpha_{bcc} = \zeta_{EP}\left(\frac{1}{2}, \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}\right) = 1.762675.
\]  

(45)

Here we have only dealt with the pure Coulomb-interaction, but this treatment can be extended to screened electrostatic interactions as well (Kanemitsu et al., 2003).

The infinite sum in (43) strongly resembles the Riemann-zeta function, except each term is weighted by a factor \((-1)^{k+1}\), and its numerical value is \(\alpha_{1D} = 2 \ln(2) \approx 1.3863\). Although in two and three dimensions the summation can be written explicitly, obtaining a precise numerical value is far from easy and the Epstein-zeta function is required. This function can be thought of as a generalised zeta-function (Ivic, 2003; Shanker, 2006) which is defined by

\[
\zeta_{EP}(s, P) = \sum_{P \neq 0} \frac{1}{P^s}.
\]  

(46)

where \(P\) is a quadratic form defined on a \(d\)-dimensional lattice. All lattice points for which \(P \equiv 0\) are excluded from the summation. This function can be analytically continued to the same domain as the Riemann-zeta function and also has its only pole at \(s = 1\) with residue \(\pi/\sqrt{\Delta}\). The similarity goes further since \(\zeta_{EP}(s, P)\) also satisfies a functional equation expressing mirror symmetry. Thus, there is an inclination to generalise Riemann’s conjecture: all non-trivial zeros of \(\zeta_{EP}(s, P)\) have real part one half. The temptation to do so is strengthened if one chooses specific quadratic forms, e.g. \(\zeta_{EP}(s, \text{Id}_1) = 2\zeta(2s)\), or \(\zeta_{EP}(s, \text{Id}_4) \sim \zeta(s)\zeta(s - 1)\), where \(\text{Id}_n\) is the \(n\)-dimensional identity matrix. Indeed, it was shown eighty years ago that for binary quadratic forms (two dimensional lattice), infinitely many zeros of \(\zeta_{EP}(s, P)\) lie on the critical line (Potter and Titchmarsh, 1935) in a similar manner to Hardy for the Riemann zeta function (Hardy, 1914). Remarkably, however, it has also been shown that in any dimension one can construct such a \(P\), that the generalised hypothesis does not hold (Terras, 1980).

This, admittedly negative, result shows the intriguing connection between crystallography and this generalised Riemann Hypothesis, but let us now depart from the abstract and static crystal structure of solids, and examine the dynamics of this system.

The lattice vibrations, phonons, are bosonic quasiparticles. Therefore if one knows their energy spectrum,
\( \hbar \omega_k \), then the total energy of the phonon gas is simply the sum over all modes of the crystal

\[
U = \sum_k \frac{\hbar \omega_k}{e^{\hbar \omega_k/k_B T} - 1}.
\]

Since the number of possible modes is large, \( 3N \), where \( N \) is the number of atoms in the lattice, one might convert this expression into an integral by introducing the phonon density of states \( g(\omega) \) normalised as \( \int g(\omega) d\omega = 3N \). Using standard methods to calculate the specific heat of the solid, a directly measurable quantity, the following expression can be obtained

\[
c_V = \int_0^\infty \left( \frac{\hbar \omega}{k_B T} \right)^2 \frac{e^{\hbar \omega/k_B T}}{(e^{\hbar \omega/k_B T} - 1)^2} \, g(\omega) d\omega.
\]

The only sample-specific quantity here is \( g(\omega) \). Surprisingly, the number theoretical Möbius function and the related Möbius inversion provides a transformation to express \( g(\omega) \) as a function of the measured specific heat

\[
g(\omega) = \frac{1}{k_B u^2} \sum_{n=1}^\infty \mu(n) \mathcal{L}^{-1} \left( \frac{c_V(h/k_B u)}{u^2} \right)
\]

where the inverse Laplace-transform, \( \mathcal{L}^{-1} \), converts the space of \( u = \hbar/k_B T \) to \( \omega/n \) (Chen, 1990). Around the same time another inversion technique appeared in the literature (Xianxi et al., 1990) and was proven to be equivalent to the one discussed above (Ming et al., 2003). In the latter formulation another special function, the Riemann zeta function was used, but in order to avoid the dependence on the unproven Riemann Hypothesis a free “regularisation” parameter \( s \) was also introduced. The density of states in this formalism is

\[
g(\omega) = \frac{1}{2\pi \omega} \int_{-\infty}^\infty \frac{\omega^{ik+1} Q(k)}{\Gamma(ik+s+2)} \zeta(ik+s+1) \, dk
\]

where \( Q(k) = \int_0^\infty u^{ik+s-1} c_V(1/u) \, du \). Physically the density of states, \( g(\omega) \) should be independent of the regularisation parameter, although the existence of \( Q(k) \) requires that \( s \) must fall into the \( 0 \leq s_1 < s < s_2 \) range where \( s_1 \) and \( s_2 \) are the exponents of the specific heat asymptotes at high- and low temperatures, respectively. Due to the Dulong-Petit law, at high temperature the specific heat is independent of the temperature, therefore \( s_1 \equiv 0 \). On the other end of the temperature scale the specific heat of phonons vanishes as \( T^d \) in \( d \) dimensions. Therefore \( \zeta(s) \) in the denominator of the integrand \( 50 \) sweeps through the \([1, 1+d]\) strip and ensures that no zeros of \( \zeta(s) \) can occur there. Summarising, the asymptotes of the specific heat contribution of lattice vibrations in a solid provide an experimentally determined zero-free region of \( \zeta(s) \) on the complex \( s \) plane. Although this offers no further restriction than that which is already known from mathematics, it is an example where physics places independent bounds upon the location of the zeros.

### E. Statistical physics

The description of both bosons and fermions relies on the mathematical properties of the Riemann zeta function. We show how the problem of the ‘grand canonical catastrophe’ of number fluctuation in an ideal Bose-Einstein condensate is connected to number theory. We introduce the concept of the primon gas, and also consider number theoretical models of Brownian motion.

Although statistical physics, the physics of systems with a large number of degrees of freedom, relied heavily upon combinatorics well before the birth of quantum mechanics, probably the first appearance of the Riemann zeta function in statistical physics occurred in Planck’s momentous work on black body radiation, the dawn of the quantum era. From then on, the Riemann zeta function pops up in numerous different branches of statistical physics, from Brownian motion to lattice gas models.

Since the topic of ultra-cold quantum gases has expanded rapidly in the past decade, we interpret the implications of the distribution of the Riemann zeta zeros in this area first. We start with the non-relativistic, non-interacting, spin zero Bose gas and treat the spatial dimension \( D \) as a free parameter. It is a standard textbook derivation (Haug, 2001) to show that this system undergoes a phase transition at low temperatures, where the de-Broglie wavelength \( \Lambda \) of the particles becomes comparable to the inter-particle distance, and thus the quantum nature of the constituents becomes decisive. Since the particles are free, their spectrum is continuous, simply equal to the kinetic energy \( \epsilon = p^2/2m \). The total number of particles \( N \) is the sum of particles in each quantum state

\[
N = \frac{1}{(2\pi \hbar)^D} \int \text{f}_{\text{BE}}(\epsilon(p)) \, d^Dp \, d^Dp \equiv \frac{V}{(2\pi \hbar)^D} \int \frac{d^Dp}{e^{(\epsilon(p)-\mu)/kT} - 1}
\]

Changing the integration from momentum to energy leads directly to

\[
N \propto \int_0^\infty \frac{e^{D/2-1}}{e^{(s-\mu)/kT} - 1} \, ds \propto \zeta \left( \frac{D}{2} \right)
\]

In the last step we used the fact that the chemical potential approaches the energy of the lowest lying state, i.e. \( \mu = 0 \).

This result shows that the Bose-Einstein condensation phase transition cannot occur in homogeneous non-interacting systems in dimensions lower than three. The total number of atoms is a positive number and fixed for our system. In one spatial dimension, since \( \zeta(1/2) < 0 \), the positivity of \( N \) cannot be fulfilled. For two-dimensions the right hand side of \( 52 \) is divergent due to the pole of the Riemann zeta function \( \zeta(s) \) at \( s = 1 \), therefore \( N \) appears to be infinite. The position of
this pole can be interpreted as the manifestation of the Mermin–Wagner–Hohenberg theorem, which guarantees that a homogeneous two-dimensional system, provided the interaction is sufficiently weak, cannot undergo a phase transition. One may thus see that the pole structure of the Riemann zeta function determines whether our system of interest can undergo a phase transition or not. We note here that this phase transition can occur in lower dimensions for inhomogeneous systems (Bagnato and Kleppner, 1991; Dai and Xie, 2003; Widom, 1968.

Let us turn to another fundamental question of statistical mechanics: the equivalency of different statistical ensembles. The difference between the predictions for the “Riemann gas” (see below) based on microcanonical, canonical, and grand-canonical ensembles has been investigated by Tran and Bhaduri (Tran and Bhaduri, 2003). Later Eckhardt extended the so-called “grand-canonical catastrophe” of an ideal Bose gas (Ziff et al., 1977). The number fluctuation of an ideal boson gas is

$$\langle \delta N \rangle^2 = \sum_{k=0}^{\infty} \langle n_k \rangle \langle n_k + 1 \rangle$$

(53)

where $\langle n_k \rangle$ denotes the ensemble average of the occupation number of the $k$th energy eigenstate. According to the formula above, in the presence of a macroscopically occupied ground state, the number fluctuation is proportional to the total number of particles, $\delta N_0 \sim N$, which, in the thermodynamical limit ($N \to \infty$), leads to divergence.

Grossmann and Holthaus examined the illustrative model system of an ideal Bose gas trapped in a $d$-dimensional potential with a power-law energy spectrum, $\epsilon_n \sim \hbar \omega n^2$, where $\nu_n$ labels the energy eigenstates (Grossmann and Holthaus, 1997b; Weiss and Wilkens, 1997). Later Eckhardt extended the analysis to the mean density of states and the level spacing distribution for ideal quantum gases (Eckhardt, 1999). Grossmann et al. showed how the dimensionality and $\sigma$, which, in some sense, measures the strength of the potential, depress or enhance the number fluctuation of the ground state as a function of the rescaled temperature $t = k_B T / \hbar \omega$:

$$\langle \delta N_0 \rangle^2 \sim \begin{cases} C t^{d/\sigma} & (0 < d/\sigma < 2) \\ t^2 \ln(t) & (d/\sigma = 2) \\ (\zeta(d/\sigma - 1)t)^2 & (2 < d/\sigma) \end{cases}$$

where $C$ is calculated from a $d$-dimensional Epstein zeta function (Holthaus et al., 2001), although here its value does not play a significant role. Therefore in a given spatial dimension the potential can enhance the fluctuation while dimensionality depresses it. They also examined the behaviour of the heat capacity around the critical temperature $t_0$ and proved that the heat capacity changes continuously at $t_0$ if $1 < d/\sigma \leq 2$, but if $d/\sigma > 2$ it undergoes a jump given by

$$\frac{C_< - C_>}{N k_B} \bigg|_{t_0} = \left( \frac{d}{\sigma} \right)^2 \frac{\zeta \left( \frac{d}{2} \right)}{\zeta \left( \frac{d}{2} - 1 \right)}$$

(54)

where $C_<$ and $C_>$ denote the asymptotic values of the heat capacity at $t \to t_0$ from below and above, respectively. It is worthwhile to note that $\langle \delta N_0 \rangle^2$ in the canonical ensemble could be expressed as the following integral over the complex plane

$$\langle \delta N_0 \rangle^2 = \frac{1}{2\pi i} \int_{\tau-i\infty}^{\tau+i\infty} \Gamma(t) \Lambda(\beta, t) \zeta(t - 1)$$

(55)

where $\Lambda(\beta, t) = \sum \langle \beta \epsilon_n \rangle^{-t}$ is the spectral zeta function of a given spectrum $\epsilon_n$, and $\tau$ is chosen so all the poles of the integrand lie on the left of the path of integration. Therefore all the results shown above are determined by the pole structure of the spectral and the Riemann zeta functions, $\Lambda(\beta, t)$ and $\zeta(s)$, respectively. The large-system behavior is extracted from the leading pole, while the finite-size corrections are encoded in the next-to-leading poles.

The formulae (54) and (55) above did not just clarify an important physical question, namely number fluctuation properties of a $d$-dimensional boson gas below the critical temperature, but also had valuable number theoretical consequences. The problem solved above is a purely combinatorial one (Grossmann and Holthaus, 1997a; Holthaus et al., Weiss et al., 2003; Weiss and Holthaus, 2002): how many ways can one distribute $n$ excitation quanta over $N$ particles? This question, for general $n$ and $N$, is quite difficult. However, in the low temperature limit the number of excitations, $n$, is much smaller than the number of particles, $N$. This problem thus becomes tractable and one could obtain the results mentioned above. Calculating the number fluctuation of a boson gas in a one-dimensional ($d = 1$) harmonic potential ($\sigma = 1$) provides $\langle \delta N_0 \rangle^2 \sim t$. But $t$ is simply proportional to the number of energy quanta ‘stored’ in the excited states, $t = (k_B T / \hbar \omega) = n$ and therefore $\langle \delta N_0 \rangle^2 \sim n$. A mathematician – according to Grossmann and Holthaus – can now interpret this formula:

If one considers all unrestricted partitions of the integer $n$ into positive, integer summands, and asks for the root-mean-square fluctuation of the number of summands, then the answer is (asymptotically) just $\sqrt{n}$.

An intriguing consequence of this analysis is that a Bose-Einstein condensate could be used (in theory at least) to factorise numbers (Weiss et al., 2004) which could be treated as a quantum computer calculating the prime factors.
Furthermore, using their physical insight, Weiss and collaborators could derive the following non-trivial number theoretical result. Let \( \Phi(n,M) \) denote the number of partitions of \( n \) into \( M \) summands regardless of their order (e.g. \( \Phi(5, 2) = 2 \) while \( \Phi(5, 4) = 1 \)), and \( \Omega(n) \) stand for the total number of different partitions, i.e. \( \Omega(n) = \sum_{m=1}^{n} \Phi(n,m) \). It is a natural step to introduce the “probability” of having exactly \( M \) terms in a random partition by \( p(n,M) = \Phi(n,M) / \Omega(n) \). It was then shown that this probability distribution does not become Gaussian, and it adopts its limiting distribution shape if \( n > 10^{10} \), which itself is a remarkable fact.

Here we only mention that the same combinatorial problem arises in many different branches of mathematical physics, such as lattice animals in statistical physics (Lima and de Menezes, 2003; Wu et al., 1996), numerical analysis on combinatorial optimisation (Andreas and Beichl, 2003; Bäuke et al., 2003; Majumdar and Krapivsky, 2002; Mertens, 1998) and also in the description of the low-energy excitations of a one dimensional fermion-system as bosonic degrees of freedom (bosonisation) (Schönhammer and Meden, 1996).

Tran and Bhaduri’s, and then Holthaus and Weiss’ works further underline that the irregular behaviour of the canonical ensemble lies in the combinatorics of partitioning integers and the microcanonical and canonical ensembles prognosticate dramatically different ground state number-fluctuations, \( \delta n_0 \). This is an important example which unequivocally shows that the standard statistical ensembles can not always be regarded as equivalent.

These examples, while not directly related to any attempt to prove the Riemann Hypothesis, but rather just the zeta function, do illustrate that results in physics can have profound implications for mathematics in general and number theory in particular.

The interpretation of prime numbers or the Riemann zeta zeros as energy eigenvalues of particles appears not just in quantum mechanics but also in statistical mechanics. Below, we review two concepts: the Riemann gas, sometimes called the primon gas, and the Riemann-Dirichlet principle in that domain (bosonisation). Julia extends the study to the grand canonical ensemble by introducing a chemical potential \( \mu \) (Julia, 1994), therefore replacing the primes \( p \) with new ‘primes’ \( pe^{-\mu} \). This generalisation of the Riemann gas is called the Beurling gas, after the Swedish mathematician Beurling who generalised the notion of prime numbers. Examining a boson primon gas with fugacity \( -1 \) shows that its partition function is just \( Z_{2s} = \zeta(2s)/\zeta(s) \).

This last result has an astonishing interpretation. We know that for a system, formed by two sub-systems not interacting with each other, the overall partition function is simply the product of the individual partition functions of the subsystems. Equation (58) has precisely this structure, there are two decoupled systems: a fermionic “ghost” Riemann gas at zero chemical potential and a boson Riemann gas with energy-levels \( E_n = 2\epsilon_0 \ln(p_n) \).

Julia also calculates the appropriate Hagedorn temperatures and analyses how the partition functions of two different number theoretical gases, the Riemann gas

The canonical ensemble is of course not the only ensemble used in statistical physics. Julia extends the study to the grand canonical ensemble by introducing a chemical potential \( \mu \) (Julia, 1994), therefore replacing the primes \( p \) with new ‘primes’ \( pe^{-\mu} \). This generalisation of the Riemann gas is called the Beurling gas, after the Swedish mathematician Beurling who generalised the notion of prime numbers. Examining a boson primon gas with fugacity \( -1 \) shows that its partition function is just \( Z_{2s} = \zeta(2s)/\zeta(s) \).

This last result has an astonishing interpretation. We know that for a system, formed by two sub-systems not interacting with each other, the overall partition function is simply the product of the individual partition functions of the subsystems. Equation (58) has precisely this structure, there are two decoupled systems: a fermionic “ghost” Riemann gas at zero chemical potential and a boson Riemann gas with energy-levels \( E_n = 2\epsilon_0 \ln(p_n) \).

Julia also calculates the appropriate Hagedorn temperatures and analyses how the partition functions of two different number theoretical gases, the Riemann gas

\[
Z_B = \sum_{n=1}^{\infty} \exp \left( -\frac{E_n}{k_B T} \right) = \sum_{n=1}^{\infty} \frac{1}{n^s} = \zeta(s)
\]

where \( s = \epsilon_0 / k_B T = \beta \epsilon_0 \) and \( \beta = (k_B T)^{-1} \) is the inverse temperature. The partition function for the primon gas is thus the Riemann zeta function \( \zeta(s) \) and hence the alternative nomenclature. It is apparent, by looking at the domain of \( \zeta(s) \), that \( Z_B \) is well-behaving for \( s > 1 \), i.e. at low-temperatures, while \( s \leq 1 \) is physically unacceptable. The boundary, \( s = 1 \) represents a critical temperature, called the Hagedorn temperature (Hagedorn, 1965) above which the system cannot be heated up, since its energy becomes infinite.

\[
\langle E \rangle = -\frac{\partial}{\partial \beta} \ln(Z_B) = -\frac{\epsilon_0}{\zeta'(\beta \epsilon_0)} \approx \frac{\epsilon_0}{s-1}.
\]
This leads to the interpretation that one of the original boson Riemann gas, which do not interact with each other.

### Mutual and Non-Mutual Interaction

The partition function of a free, non-interacting $\kappa$-parafermion gas can be shown to be (Bakas and Bowick, 1991)

$$Z_n(s) = \frac{\zeta(s)}{\zeta(k s)}.$$  \hfill (59)

Bakas further demonstrates, using the Dirichlet convolution ($\ast$), how one can introduce free mixing of parafermions with different orders which do not interact with each other

$$f \ast g = \sum_{d|n} f(d) g\left(\frac{n}{d}\right).$$  \hfill (60)

where the shorthand notation $d|n$ means $d$ is a divisor of $n$. This operation preserves the multiplicative property of the classically defined partition functions: $Z_{\kappa_1 \ast \kappa_2} = Z_{\kappa_1} Z_{\kappa_2}$. It is even more intriguing how interaction can be incorporated into the mixing by modifying the Dirichlet convolution with a kernel function or twisting factor

$$f \ast g = \sum_{d|n} f(d) g\left(\frac{n}{d}\right) K(n, d).$$  \hfill (61)

Using the unitary convolution Bakas establishes a pedagogically illuminating case, the mixing of two identical boson Riemann gases. He shows that

$$(Z_\infty \ast Z_\infty) = \frac{\zeta^2(s)}{\zeta(2s)} = \frac{\zeta(s)}{\zeta(2s)} \zeta(s) = Z_2 Z_\infty.$$  \hfill (62)

Thus mixing two identical boson Riemann gases interacting with each other through the unitary twisting, is equivalent to mixing a fermion Riemann gas with a boson Riemann gas which do not interact with each other. This leads to the interpretation that one of the original boson components suffers a transmutation into a fermion gas. It is noteworthy to mention that the Möbius function, which is the identity function with respect to the $\ast$ operation (i.e. free mixing) reappears in supersymmetric quantum field theories as a possible representation of the $(-1)^F$ operator, where $F$ is the fermion number operator (Spector, 1989, 1990, 1998). In this context, the fact that $\mu(n) = 0$ for square-free numbers is the manifestation of the Pauli exclusion principle.

It is therefore interesting that, what initiated as rather academic studies to investigate potential attacks on the Riemann Hypothesis, may lead to advances in physics. But let us return to the Hypothesis through a slightly different definition of the Riemann gas. Here the energy of the ground state is taken to be zero and the energy spectrum of the excited state is $\epsilon_n = \ln(p_n)$, where $p_n$ (n = 2, 3, 5, ...) runs over the prime numbers. Let $N$ and $E$ denote the number of particles in the ground state and the total energy of the system, respectively. As we demonstrated above, the fundamental theorem of arithmetic allows only one excited state configuration for a given $E = \ln(n)$ (n is an integer). It immediately means that this gas preserves its quantum nature at any temperature, since only one quantum state is permitted to be occupied. The number fluctuation of any state (the ground state included) is therefore zero. In contrast, the $\delta n_0$ predicted by the canonical ensemble is a smooth non-vanishing function of the temperature, while the grand-canonical ensemble still exhibits a divergence. This discrepancy between the microcanonical (combinatorial) and the other two ensembles remains even in the thermodynamic limit.

One may argue that the Riemann gas is fictitious and its spectrum is unrealisable. However, the spectrum $\epsilon_n = \ln(n)$ does not increase with $n$ more rapidly than $n^2$, therefore the existence of a quantum mechanical potential supporting this spectrum is possible (cf. inverse scattering transform used in section III.B). The potential has been given in (Weiss et al., 2004):

$$V(x) = V_0 \ln \left(\frac{|x|}{L}\right).$$  \hfill (63)

where $V_0$ and $L$ are positive constants. Within the semiclassical approximation the spectrum of this potential is

$$\epsilon_n = V_0 \ln(2n + 1) + V_0 \ln \left(\frac{h}{2\pi \sqrt{2mV_0}}\right)$$  \hfill (64)

where $n = 0, 1, \ldots$ and the second term only represents a constant energy shift.

Recently, LeClair published two works (LeClair, 2007, 2008) developing and applying a finite-temperature field theoretical formalism for both boson and fermion gases in low spatial dimensions in which he efficiently disentangles zero temperature dynamics and quantum statistical sums for both the relativistic and non-relativistic cases. His alternative approach is based on an $S$-matrix formulation of statistical mechanics (Dashen et al., 1964), which
redefines the quantum statistical mechanics directly in terms of dynamical filling fractions, \( f(k) \). Assuming the two-body scattering kernel, \( K \), is constant (i.e. constant scattering length) he derives, as pedagogical examples, the well-known results for the boson

\[
T_c \sim \left( \frac{n}{\zeta(d/2)} \right)^{2/d}
\]

and also for the fermion gas

\[
\varepsilon_F \sim \left[ \Gamma \left( \frac{d+2}{2} \right) n \right]^{2/d}.
\]

In two dimensions the critical temperature for the boson gas vanishes because of the \( \zeta(s) \) divergence at \( s = 1 \), therefore this dimension needs further consideration. Due to this instability, LeClair extends the examination for energy-dependent two-body kernels, \( K = -\Re(\gamma_k^{2\nu-1}) \) \( (\nu \) is a complex number and \( \gamma_k \) is constant), for a one-dimensional fermion gas and explicitly constructs a quasi-periodic potential, \( V(x) \sim \cos(\log(x)/x^{2\nu}) \), in the real space which reproduces the given kernel \( K \) in the two-body scattering approximation. Furthermore, the thermodynamic variables, such as density and pressure, are also shown to be physically valid (i.e. positive and have finite value) provided \( 1/2 < \Re(\nu) < 3/2 \). This fully covers the right hand side of the critical strip divided by the critical line, and due to the symmetry of \( \zeta(s) \) this half-strip can be extended to the whole critical strip. His argumentation is based on both the non-vanishing, non-divergent nature of the physical quantities and also on the assumption that an interaction necessarily modifies the thermodynamical quantities. If \( \zeta(\nu) \) would be zero somewhere in the critical strip, but off the critical line, then the leading order contribution to the thermodynamical quantities would not be zero contradicting the original assumption – LeClair argues. This contradiction led him to conclude that \( \zeta(\nu) \) must be non-zero in the \( 1/2 < \nu < 3/2 \) strip, which can automatically be extended to the whole critical strip by using the symmetries of the Riemann zeta function. LeClair, therefore, claims: \( \zeta(\nu) \) can have no zeros in the given range, consequently the Riemann Hypothesis must be true. The basis for this conclusion however, is itself an assumption and so does not constitute a proof of the Riemann Hypothesis, but does provide another point of attack.

Examination of a similar fictitious, fermionic, many-body system has also been considered by Leboeuf and lead to the conclusion that “time-periodic dynamical evolutions have to be considered as serious candidates for the Hilbert-Pólya Hamiltonian” [Leboeuf et al., 2001].

At the end of this section, let us mention an interesting interlocking area of statistical physics and number theory. A few authors have focused on the connection between number theoretical functions and Brownian motion [Billingsley, 1973; Evangelou and Katsanos, 2005; Good and Churchhouse, 1968; Shlesinger, 1986; Wolf, 1998] or percolation [Varel, 1998]. The connection seems to be suggestive, especially if one defines the random motion through the Möbius function \( \mu(n) \), i.e. if \( \mu(n) = \pm 1 \) the particle moves up- or downwards, and if \( \mu(n) = 0 \) it does not move. Therefore the distance of the particle from the origin after \( n \) steps is \( M(n) = \sum_k \mu(k) \). The importance of this kind of Brownian motion lies in the so-called Mertens conjecture. This states if \( |M(n)| \leq \sqrt{n} \) then the Riemann Hypothesis is true.

Figure 9 shows the path of the particle for the first million steps. Although it is not possible to conclude: the cumulative sum of \( \mu(n) \) remains bounded by \( \pm \sqrt{n} \), this conjecture would actually be wrong as te Riele and Odlyzko indirectly proved [Titchmarsh and Heath-Brown, 2003]. There is no explicit counterexample known, but we have a loose interval \( [10^{14}, \sim 3.6 \times 10^{10^{57}}] \) in which there exists an \( n \) such that \( M(n)/\sqrt{n} > 1 \) [Kotnik and Riele, 2006]. Nevertheless, the Mertens conjecture is a sufficient condition for the Riemann Hypothesis to be true, not a necessary one. Its falsity therefore cannot invalidate the Riemann Hypothesis. The failure of the Mertens Conjecture at such a high \( n \) value, however, does give cause for concern regarding numerical evidence for the validity of the Riemann Hypothesis.

However, this is not the only possibility to define a random walk either on the \( \zeta(s) \) zeros or on the prime numbers. In the early 1970’s Billingsley defined a random, but finite, walk [Billingsley, 1973] based on the fundamental theorem of arithmetic.

Let \( f(n) \) denote the number of prime factors of \( n \) not counting their multiplicity, e.g. \( f(40) = 2 \), since \( 40 = 2^3 \times 5 \). It can be shown, that on average, numbers below \( N \) have \( \ln(\ln(N)) \) factors; a result which on its own is a surprise. For example, numbers below \( e^{10} \approx 10^{59} \) have only 10 distinct factors on average. Based on the
factorisation one can define the following random walk: choose an integer \( n \in [0, N] \), starting from the origin we go up by a unit if 2 divides \( n \) and down if it does not, and continue the test with 3, 4, \ldots. Although this construction does not seem to be as random as a coin-tossing random walk and has few flaws (e.g. it is biased), Billingsley suggests a remedy to these problems and shows how the similarity to Brownian motion leads to an Erdős–Kac central limit theorem for \( f(n) \)

\[
P\left( \alpha \leq \frac{f(n) - \ln(\ln(N))}{\sqrt{\ln(\ln(N))}} \leq \beta \right) \to \frac{1}{\sqrt{2\pi}} \int_{\alpha}^{\beta} e^{-u^2/2} \, du.
\]  

(67)

Therefore, the probability of \( f(n) \) not deviating from the expected value \( \ln(\ln(N)) \) more than \( \alpha \) or \( \beta \) times the standard deviation can be estimated by a Gaussian integral. Therefore, the mapping of the number theoretical problem onto a Brownian motion helps to derive a limit theorem for the number theoretical function \( f(n) \). As an example, if we chose \( \alpha = -1 \) and \( \beta = 1 \) for \( N = 10^{8566} \) gives \( P(-1 \leq (f(n) - 10)/\sqrt{70} \leq 1) \approx 0.68 \), thus approximately 70% of the numbers below the chosen \( N \) have from 6 to 13 distinct prime factors.

M. Wolf defined random walks in a different way [Wolf, 1998] and could examine the distribution and correlation of twin-primes (where \( p \) and \( p + 2 \) are both primes) and also of cousin primes (\( p \) and \( p + 4 \) are both primes). He also suggested new random number generators with theoretically infinite period based on this kind of random walk, contrary to the widely used random number generators [Press et al. 2007]. He also argues with and with computations demonstrates the multi-fractal nature of a subset of prime numbers [Wolf, 1989].

IV. CONCLUSION

‘All results of the profoundest mathematical investigation must ultimately be expressible in the simple form of properties of the integers.’

(Leopold Kronecker)

Since this review is a summary itself in some sense, here we only attempt to conclude with some general remarks.

In many respect the history of the Riemann Hypothesis is very similar to that of Fermat’s Last Theorem, which was stated in the seventeenth century and solved 358 years later [Aczel, 1997; Ribenboim, 1999], and along the path towards the final proof it inspired and gave birth to new areas of mathematics, such as the theory of elliptic curves. Although the Riemann Hypothesis has not been proven or disproven it has already stimulated and influenced many areas of mathematics, e.g. \( L \)-functions, which can be thought of as generalised zeta functions, for which a generalised Riemann Hypothesis may hold. Interestingly, for \( L \)-functions defined over functional space rather than the number field, the similar hypothesis is rigorously proven.

Further evidence also suggests the validity of the Riemann Hypothesis, let us just think of Levinson’s theorem guaranteeing that at least one third of the zeros are on the critical line. We, however, cannot exclude the possibility of the existence of a counterexample to the Riemann Hypothesis, i.e. a very high lying zero \( s = \sigma + it \) for which \( \sigma \neq 1/2 \). Similarly to the Mertens conjecture, the counterexample may occur so high on the critical line, that we have no machinery to even calculate zeros at that elevation. The immediate impact of such a collapse of the Riemann Hypothesis would be immense since there exist numerous “proofs” that are contingent upon it [Titchmarsh and Heath-Brown, 2003].

That said, we cannot miss out in this review one computational masterpiece. Not long after World War II, in which mechanical and electrical ‘computers’ were often used for encrypting messages (Enigma) and also for research (ENIAC), Balthasar van der Pol constructed an electro-mechanical machine which could calculate the first few zeros of the Riemann zeta function [van der Pol, 1947]. This construction, despite its limited achievement, deserves to be treated as a gem in the history of the natural sciences. Several decades later, on the other end of the spectrum, a state-of-the-art application of numerical techniques carried out by Brent, van de Lune, te Riele and Winter [Brent, 1979; Brent et al., 1982; van de Lune and te Riele, 1983; van de Lune et al., 1986] calculated the first \( 1.5 \times 10^9 \) zeros. Meanwhile Odlyzko [Odlyzko, 1987] explored the zeros located around \( t \sim 10^{20} \), and showed that all zeros (millions of them) he found do exactly lie on the critical line. Here we note, that these numerical checking are, of their own right, significant achievements, and also have influenced the development of fast numerical techniques used in physics (see e.g. [Draghicescu, 1994; Greengard, 1994]).

In this review article we collected a few examples from different areas of mathematical physics, starting with classical mechanics and finishing with statistical mechanics, where the Riemann zeta function \( \zeta(s) \), especially its zero- and pole-structure, has a highly influential role.

In the section devoted to classical mechanics, we showed how the Riemann Hypothesis can arise in a simple mechanical system, a ball bouncing on a rigid wall. We also argued how these billiard systems lead to a revolutionary new way of describing the dynamics of a chaotic system by introducing the evolutionary operator. Here we also sketched the connection, a trace formula, between the dynamics of a chaotic system and the periodic orbits of the same system.

This new descriptive language of dynamics through the trace formulae of the Green function is suitable to develop a new quantisation technique for chaotic quantum systems which otherwise was impossible using the stan-
standard Bohr quantisation rules. Gutzwiller’s trace formula has been explicitly mentioned, because the Riemann zeta function obeys a very similar expression. Therefore, we could compare the two formulae, \( \zeta(s) \) and \( \phi_n(s) \), and imagine what properties a quantum system might have if its spectrum mimicked the zeros of the Riemann zeta function.

We also surveyed two other attempts to find a quantum system which has a connection to the Riemann zeta function. Both of these directions try to associate \( \zeta(s) \) with the spectrum of the system. The difference between these approaches is that one of them relates \( \zeta(s) \) to the positive energy spectrum, i.e. scattering states, while the other, based on the Hilbert-Pólya conjecture, proposes systems where the negative energies, thus the bound states of the system, coincide with the zeros of \( \zeta(s) \). This latter case naturally guides us to condensed matter physics and statistical mechanics, where one has to evaluate physical observable on the lattice points, or derive all thermodynamical properties of a given particle-system provided the spectrum is given.

In the sections concentrating on condensed matter physics, we first showed how the Riemann zeta function, or one of its ancillary functions, arose when we calculated the binding energy of a given structure of solid matter. Finally we showed how physical requirements for the specific heat of a solid can provide zero-free regions for the Riemann zeta function. Research in this direction eventually may offer narrower zero-free regions, and complement the approach in pure mathematics.

In the last section, we discussed three main areas of statistical physics where the Riemann zeta function and its number theoretical aspects influence the behaviour of a physical system. Firstly, we considered the low-temperature phase transition of bosons, and showed that the pole structure of \( \zeta(s) \) prohibits Bose-Einstein condensation in one- and two-dimensional uniform systems. We also reviewed the ‘grand canonical catastrophe’ of an ideal Bose gas, where the predictions of two ensembles widely used in statistical physics contradict each other, showing, therefore, that these ensembles cannot be equivalent to one another. Finally, we examined a possible Brownian motion model for the number theoretical Möbius function.

It would not be without precedent if a completely new theory or a new mathematical language is needed in which the Riemann Hypothesis can be ‘worded’ naturally for the hypothesis to be finally proved. As has happened earlier with mathematics, natural science, and in particular physics, can give impetus and motivate new directions perhaps leading to the final proof. It is amazing and captivating to see that a purely number theoretical function has so many direct links to classical and modern physics.

Nowadays we are not surprised by Galileo’s famous keynote: “[Nature] is written in the language of mathematics, and its characters are triangles, circles, and other geometric figures without which it is humanly impossible to understand a single word of it; without these, one wanders about in a dark labyrinth” (Drake 1957). Probably we are not meandering in a labyrinth, but we are definitely puzzled by the overwhelming difficulty of proving the Riemann Hypothesis. We simply do not know as yet whether physics will ultimately help in understanding such an elegant mathematical statement as the Riemann Hypothesis, but we are definitely witnessing the intertwining and invigoration of both disciplines. The authors can only express their hope that this work has to some extent captured the imagination of the Reader and, if so, it has fulfilled its intended aim.

ACKNOWLEDGMENTS

Daniel Schumayer expresses his gratitude to Brandon P. van Zyl for creating an inspiring and welcoming environment and partial financial support through a grant from the National Science and Engineering Research Council of Canada. We would also like to sincerely thank all the referees of this manuscript. Their comments and suggestions were highly valuable and lead to a significant increase in the final quality of this paper.

This work has been financially supported by the University of Otago and the Government of New Zealand through the Foundation for Research, Science and Technology under New Economy Research Fund Contract No. NERF-UOOX0703.

REFERENCES

Ablowitz, M. J., and P. A. Clarkson, 1991, Solitons, nonlinear evolution equations and inverse scattering, volume 149 of London Mathematical Society Lecture Note Series (Cambridge University Press).

Aczel, A. D., 1997, Fermat’s Last Theorem: Unlocking the Secret of an Ancient Mathematical Problem (Dell Publishing).

Agam, O., B. L. Altshuler, and A. V. Andreev, 1995, Phys. Rev. Lett. 75(24), 4389.

Alfaro, V. D., and T. Regge, 1965, Potential scattering (North-Holland, Amsterdam).

Alt, H., C. I. Barbosa, H. D. Gräf, T. Guhr, H. L. Harney, R. Hofferbert, H. Rehfeld, A. Richter, 1998, Phys. Rev. Lett. 81(22), 4847–4850.

Andreas, A. K., and I. Beichl, 2003, Computing in Science & Engineering 5(1), 48.

Aneva, B., 1999, Phys. Lett. B 450(4), 388.

Aneva, B., 2001a, Disordered and Complex Systems 553, 173.

Aneva, B., 2001b, AIP Conf. Proc. 553(1), 173.

Arnol’d, V. I., 1997, Mathematical methods of classical mechanics, volume 60 of Graduate Texts in Mathematics (Springer), 4th edition, 978-0-387-96890-2.

Bäcker, A., R. Schubert, and P. Stifter, 1998, Phys. Rev. E 57(5), 5425.

Baeriswyl, D., D. K. Campbell, J. M. Carmelo, F. Guinea, and E. Louis (eds.), 1995, The Hubbard Model: Its Physics
Phys. Rev. E 68(2), 026206.
Schaden, M., 2006, Phys. Rev. A 73(4), 042102.
Schnielman, A. I., 1974, Usp. Mat. Nauk. 29, 181.
Schönhammer, K., and V. Meden, 1996, Am. J. Phys. 64(9), 1168.
Schumayer, D., B. P. van Zyl, and D. A. W. Hutchinson, 2008, Phys. Rev. E 78(5), 056215.
Selberg, A., 1949, The Annals of Mathematics 50(2), 297.
Shanker, O., 2006, J. Phys. A 39(45), 13983.
Sierra, G., 2007, N. Phys. B 776(3), 327.
Sierra, G., 2008, New J. Phys. 10(10), 033016.
Spector, D., 1989, Phys. Lett. A 140(6), 311.
Spector, D., 1990, Commun. Math. Phys. 127(2), 239.
Spector, D., 1998, J. Math. Phys. 39(4), 1919.
Stöckmann, H. J., and J. Stein, 1990, Phys. Rev. Lett. 64(19), 2215.
Stöckmann, H.-J., 1999, Quantum Chaos: An Introduction (Cambridge University Press).
Terras, A., 1980, J. Num. Theo. 12, 258.
Titchmarsh, E. C., and D. R. Heath-Brown, 2003, The Theory of the Riemann Zeta-function (Oxford University Press), 2nd edition.
Toet, D., and M Potemski, Y. Y. Wang, J. C. Maan, L. Tapfer, K. Ploog, 1991, Phys. Rev. Lett., 66(16), 2128.
Tomsovic, S., and E. J. Heller, 1991, Phys. Rev. Lett. 67(6), 664.
Tran, M. N., and R. K. Bhaduri, 2003, Phys. Rev. E 68(2), 026105.
Twamley, J., and G. J. Milburn, 2006, New J. Phys. 8, de la Vallée-Poussin, C. J., 1896, Ann. Soc. Sci. Bruxelles 20, 183.
Vardi, I., 1998, Experimental Mathematics 7(3), 275.
de Verdière, Y. C., 1985, Commun. Math. Phys. 102, 497.
Waldschmidt, M., P. Moussa, J. M. Luck, and C. Itzykson (eds.), 1995, From Number Theory to Physics (Springer-Verlag Berlin and Heidelberg GmbH & Co. K), 2nd edition.
Wardlaw, D. M., and W. Jaworski, 1989, J. Phys. A 22(17), 3561.
Weidenmüller, H. A., and G. E. Mitchell, 2009, Rev. Mod. Phys. 81(2), 539.
Weiss, C., M. Block, M. Holthaus, and G. Schmieder, 2003, J. Phys. A 36(7), 1827.
Weiss, C., and M. Holthaus, 2002, Europhys. Lett. 59(4), 486.
Weiss, C., S. Page, and M. Holthaus, 2004, Physica A 341, 586.
Weiss, C., and M. Wilkens, 1997, Optics Express 1(10), 272.
Widom, A., 1968, Phys. Rev. 176(1), 254.
Wigner, E., 1934, Phys. Rev. 46(11), 1002.
Wigner, E. P., 1951, Math. Proc. Cambridge Philos. Soc. 47, 790.
Wolf, M., 1989, Physica A 160(1), 24.
Wolf, M., 1998, Physica A 250(1-4), 335.
Wu, F. Y., G. Rollet, H. Y. Huang, J. M. Maillard, C.-K. Hu, and C.-N. Chen, 1996, Phys. Rev. Lett. 76(2), 173.
Wu, H., and D. W. L. Sprung, 1993, Phys. Rev. E 48(4), 2595.
Wu, H., and D. W. L. Sprung, 1995, Phys. Rev. E 51(6), 6327.
Xianxi, D., X. Xinwen, and D. Jiqiong, 1990, Beijing International Conference on High Temperature Superconductivity 521.
Zelditch, S., and M. Zworski, 1996, Commun. Math. Phys. 175(3), 673.
Ziff, R. M., G. E. Uhlenbeck, and M. Kac, 1977, Phys. Rep. 32(4), 169.
van Zyl, B. P., and D. A. W. Hutchinson, 2003, Phys. Rev. E 67(6), 066211.