Symmetry argument for the Riemann hypothesis, universality & broken symmetry

by Scott Allen[†]

Abstract:

Upon careful re-examination of Riemann's original work in the analytic continuation of the function $\varsigma(s)$ throughout the complex plane in general, and across the critical strip in particular, a natural invariant-symmetry principle and identity is derived, showing the hallmark characteristics of *universality* in the asymptotic limit. This invariant identity can be compared to Pitkänen and Castro, Mahecha et al. through the non-orthogonal coherent states, which have intriguing connections to the so-called Berry phase or Wess-Zumino term as described through an *action* in the non-linear sigma model revealing potential consequences to Riemann's Hypothesis (RH). While theirs' and others' work is primarily based on the Hilbert & Polya's conjecture as represented in supersymmetric quantum theory, conformal and scale invariance, 1/f noise, etc., the approach taken here *is not* based on any particular formalism as applied to operator or spectral theory; nor does it depend on analogues of analytic number theory to chaos theory or physics in general, requiring little more than basic complex analysis and Cauchy's Residue theorem pertaining to the zeta function itself.

[†] Comments welcome: scott@zetabase.com

I. Introduction

In recent years there have been significant advances in establishing a deeper analogy between the mathematics that underlie zeroes of the zeta function and seemingly *universal* characteristics of *chaotic* dynamic phenomena, associated with so many apparently unrelated physical systems. One of these characteristics is that any system, whatever it may comprise, in whatever form it may manifest and whatever external field or environment it couples to, will under certain conditions invariably respond in precisely the same manner as any other otherwise apparently unrelated system, under similar conditions but obviously in very different contexts. This idea underscores a very basic (non-descript) notion of *universality* as inspired by some measuring device or experience of some entity sensing a crossover between two generic states (or *phases*) of *being* or *becoming*. While the mathematics side has been the crucible and framework of this idea, involving so many areas in arithmetic, number theory, analysis and of course the Riemann Hypothesis (RH) [1-19], the physics side has continued to be an inspiration for its application in so many fields of endeavour involving very specialised expertise in nuclear / high energy physics [20-23] and scattering theory [24-26], condensed matter physics [27-32], quantum and statistical field theory [33-35], classical and quantum chaos [35-38], self-organised criticality [35-36,39],[1] and cosmology [42-47].

To illustrate with (slightly) more context and detail, consider a particular physical system of specified composition, size and dimension. If one were to cool it down, then under certain conditions the response of that system, as measured by the susceptibility of, say a magnet or compressibility of a fluid, may tend to diverge in a scale invariant manner, indicating an extremely steep yet smooth variation from one state (or *phase*) of matter to another. Provided the model[2] is set up properly, the dynamical linear response must be *unique* and evolve (piecewise) continuously in a classical or quantum sense, crossing over from an initially diffusive, apparently “random” phase to one that is “spontaneously” ordered.[3] Evidence continually has shown that this crossover is well represented by spectral eigenstates whose eigenvalues manifest as zeroes of the zeta function itself [1,37]. The *initial symmetric* phase is usually but not always modelled under the classical theory of diffusion or dissipative phenomena,[4] appropriate to those systems represented by stochastic differential equations [48]. The latter *less*-symmetric phase is usually more predictable and more *evolved* as revealed in the settling down of global *coherent* long-period orbits [49] propagating in time [37], or equivalently though a Wick rotation manifesting in a frozen topological [50] crystalline or amorphous *quasi*-ordered phase in Euclidean space. In this context, the traditional correspondence of the classical as *deterministic* and the quantum as *random*, respectively may not always apply. Indeed, the reverse may be true under certain conditions as is seen in classical diffusion vs. quantum entanglement [66-67], cosmological evolution [68] and in many inherently noisy self-interacting systems exhibiting feedback and phase locking [33,51-52].

---

[1] The idea of self-organised criticality and 1/f noise has been tied to biology [40-41] and even the concept of *primary* evolution [see M. Watkins website: http://secamlocal.ex.ac.uk/~mwatkins/isoc/evolutionnotes.htm]
[2] The model is often set up as a *Hamiltonian,* which includes all non-*universal* parameters like interactions, external field, characteristic length, lattice constant and physical constants, etc.
[3] Some researchers like to put the prefix *quasi*- in front of *random*, *ordered* or even *deterministic*, as each term is somewhat idealistic, subjective and perhaps effectively wrong when taken out of context.
[4] This includes the Langevin equation and the Fokker-Planck equation, which has applications to the pricing of options using the no-arbitrage or risk-neutral gauges under the assumptions of (market) closure.

It is commonly known that Hilbert & Polya uncovered a deep connection between Riemann Hypothesis and quantum physics, which simply states that "if the zeros of the zeta function can be interpreted as eigenvalues of $\frac{1}{2}+i\cdot\hat{T}$, where $\hat{T}$ is a self-adjoint (hermitian) operator on some Hilbert space, and since the eigenvalues of $\hat{T}$ are real, then RH follows".[5] This insight is widely considered to be the most promising approach towards establishing once and for all the truth of RH; and yet, no consistently quantised operator representing a precise physical process encompassing a *universal*, *unique* and *ubiquitous* potential of a (Schrödinger) Hamiltonian has been found.[6] While the Hilbert-Polya conjecture and thus RH remain open problems, some important influential early steps have been taken towards uncovering a deeper connection between these two conjectures, forging an even stronger link between mathematics and physics. Of particular note is the work done by Selberg [53] who came up with the famous trace formula and Gutzwiller [49,54-55] who, with help from Poincaré and Einstein, extended Selberg's work and put classical and quantum chaos on a firmer foundation through the more established theories of ergodicity and periodic orbits. Montgomery & Dyson [18,20] made a connection of particular importance between the zeroes and random quantum matrices applicable to heavy nuclei [20-22] and the GUE, not to mention the impressive strides made by Odlyzko in determining the distribution of zeroes numerically [19], confirming Montgomery's correlation of zeros [18] on the line and verification of RH to heights within the critical strip that approach Avogadro's number.

Berry & Keating [37] have conjectured that the Hilbert-Polya operator may be expressed classically as a Hamiltonian $H = x\,p$.[7] They further clarified the remaining open issues that must be resolved if one is to find a unique quantised Hermitian operator, and in so doing establish the truth of RH once and for all. In particular, they provided a strong connection between quantum chaos and the so-called Riemann (p-adic, adelic) dynamics through an asymptotically rigorous analogy by isolating certain *universal* characteristics of the periodic orbits through use of a universal sum rule [56] that is well represented by the Selberg trace formula. This analogy applies over a broad range of parameters for any physical system belonging to the same universality class, at all wavelengths or frequencies between two cut-offs specified by an ultraviolet (UV) high frequency / short wavelength limit and an infrared low frequency / long wavelength (IR) limit where dynamics manifest as periodic orbits in the *semiclassical* limit – i.e. as $\hbar \to 0$. This is commonly associated with the *mesoscopic* regime as described above, whose method of approach [37, 56] is naturally derived from Liousville's Theorem but is not considered an ensemble per se. A key characteristic here is evidence of broken time reversal symmetry, indicating that the system is ergodic, evolving steeply yet smoothly between the classical to the quantum state of *being* or *becoming*. From another point of view the system exists in a state, which is always between but never ideally *random* (stochastic) or *deterministic*. Recently, a striking connection between these *ideals* has been uncovered in higher dimensional cosmologies on non-compact manifolds, such that both may interpret different aspects of the same physical system or universe [69] showing similar characteristics to entanglement [66].

---

[5] Quote attributed to Daniel Bump – for more see http://secamlocal.ex.ac.uk/~mwatkins/zeta/physics1.htm.
[6] Exactly when this apparently *universal* connection to physics was first recognised is not entirely clear but probably goes all the way back to Riemann himself – see Marcus de Sautoy's [71] and John Derbyshire's [72] accounts of Keating's view of the *Nachlass*.
[7] This Hamiltonian has been difficult to consistently quantise and isolate physically, yet the mathematical framework is already established by Connes in the non-commutative geometry – see e.g. ref. [70].

It appears the Quantum – Riemann Dynamics analogy [37] may also apply to systems under the most stringent criteria imaginable, viz. through some all-encompassing theory between the very small (as applied through quantum theory) and the extremely large (as applied in gravitation and cosmological theory), where the UV limit in *reciprocal* space-time is now the order of the Planck length, while the IR limit is the radius of the *observable* universe [42-47]. Whichever specific or general context applies, the semi-classical or *mesoscopic* representative approach appears to be *universal* in a formal and measurable sense, showing *gauge invariance* thus is not specifically dependent on the non-*universal* nature or choice of gauge describing the potential of the Hamiltonian. Based on some recent novel approaches [57-59] the answer shows evidence it may arise *endogenously* in the zeta function itself, tying in RH directly to all candidate Hamiltonian operators obeying, say *super-symmetry* in quantum and conformal field theory or to 1/f noise involving phase modulation and phase locking [51,52] – see also [42].

A common theme throughout the works cited above is the concept of *coherent* states, which can be generated in the second quantisation as eigenvectors of a creation and / or annihilation operator describing a continuous quantum *phase*. These coherent states were picked up in the work of Pitkänen [57] and Castro & Mahecha et al. [59], who applied supersymmetric conformal invariance and quantum theory as a strategy in an attempt to prove RH. Although the work of Wolf [33] and Planat et al. [51-52] is not related to theirs in terms of application, the commonality of coherent states and its tie in to RH is truly amazing; especially considering the work of Berry [50] on adiabatic changes of an external magnetic field on quantum spins, implying the Aharonov-Bohm effect as observed in low temperature and low dimensional condensed matter systems [27-31]. Common to all these systems is their tie in to the topological effect of the Wess-Zumino term representing the quantum states that, under certain conditions, show chaos as brought about through energy states that are nothing but Riemann's zeros on the critical line implying the truth of RH, at least within the specified range. This is described *universally* through the Lagrangian or *action* by the non-linear sigma model (NLσM) – otherwise known as the Wess-Zumino-Witten model of conformal theory and is so universal that it appears to apply equally well to the metric of cosmological (Einstein) field theory in higher dimensions through the Killing norm [43,61] – for survey of recent results connecting black holes, fractals, CFT and RH see [42].

The work presented here contains very little formalism other than the basics of complex analysis and Cauchy's Integral Theorem of residues known at the time of Riemann. It is not reliant on Hilbert & Polya's conjecture or any analogy associated with the chaos of physics, biology, sociological systems or the mathematics of number theory; however, highlighted here are those aspects deemed relevant to a few of these applications. Section II provides the reader with an outline of the approach, motivation, summary and framework for this paper, while section III gives necessary background in context of the relevant work in the earlier half of the last century and a half. Sections IV and V comprise the main part of the paper, providing a formulation of the problem and isolating precisely consequences of assuming RH is true or otherwise. Appendices I, and II provide simple examples to illustrate the basic concepts, issues and support the main sections regarding definition and single-valued nature of the *argument* or phase of an arbitrary analytic function, smoothness criteria and consequences of exogenously imposing zeros off the critical line.

II. Approach and Summary

Presented here is a possible explanation for how and why Riemann evidently believed his conjecture RH was true. A concerted attempt is made to isolate and expose the problem in the most simple manner possible, by basically comparing the analytic and symmetric properties of the holomorphic function uncovered by Riemann in 1859 [1] – see Eq.(1). The approach closely follows his founding work in the hope of unravelling a short cut to a fundamental proof. It is the author's belief that axioms and theorems of complex analysis known at that time are enough to establish the truth of RH. The basic idea is to take a naïve approach and feign one has no more knowledge than what was available at the time Riemann published his classic paper [1], in the hope of gaining some insight into his motivation just as a detective would try to unravel some virtuous crime.[8] The hope is to uncover a *universal* invariant for the argument of the zeta function at points T higher up across the critical strip and thus determine how and why RH is a necessary and sufficient condition for the symmetric, single-valued and (piecewise) continuous nature of the zeta function and its variants.

In its original incantation $\varsigma(s) = \sum_{n=1}^{\infty} n^{-s}$ is convergent in the half plane, $\mathrm{Re}(s) > 1$. Riemann applied analytic continuation to show that it can also be well defined and expressible in various representations over the entire complex plane except at the single point $s = 1$. In this paper extensive reference is made to the symmetric and complex analytic properties of the holomorphic function:

$$\xi(s) := s(s-1)\cdot\frac{\varsigma(s)}{\Omega(s)} = \xi(1-s), \qquad \Omega(s) = \frac{2\pi^{s/2}}{\Gamma(s/2)}. \tag{1}$$

Here $\Omega(s)$ can be thought of as the total solid angle in s complex dimensions.

The basic strategy is to investigate consequences of a counterexample to RH using an adjustable closed rectangular contour symmetrically placed across the critical line $\mathrm{Re}\, s = 1/2$ at a specified width, $1 + 2\cdot\varepsilon$ and height T above the real axis – see Figure 1. Arguments are made in context of where zeros are located relative to the placement of the contour given by the parameter ε, which is adjustable and lies anywhere between $-1/2$ and ∞, provided no part of the contour lies on the critical line, a pole, or equivalently a zero. Through use of Cauchy's residue theorem this encompasses any and all zeros enclosed by the contour. In Riemann's original paper he chose the contour to lie along the edge of the critical strip, viz. $\varepsilon = 0$.[9] For his analysis, which fleshes out Riemann's approach, Backlund [17] chose $\varepsilon = 1/2$.

---

[8] This is not to ignore the great mathematical advances that have occurred since Riemann and references are made where appropriate, viz. the Prime Number theorem – see next footnote – and the subsequent work of Hardy and Littlewood [2-4], Bohr-Landau [5], Selberg [6, 7], Turing [8], Kac, Erdös, Weil, etc. etc.; and, in particular Siegel's work [9, 10] in deciphering Riemann's handwritten notes. See Edwards [16] for more details.

[9] This implies he believed or could prove no zeros reside on the line $\mathrm{Re}(s) = 1$ – the actual proof of which, as a consequence of the Prime Number Theorem, was provided independently years later, with help from von Mangoldt [11], by Hadamard [12] and de la Vallée Poussin [13], and much later still, using elementary real analysis by Erdös and Selberg [7].

By symmetry of the real and imaginary components of $\log \xi(s)$ across the critical strip and along the critical line one need only necessarily take half the original closed contour if we are to concern ourselves with just the imaginary components of the integral – see Figure 1 and comparing (7) with (15). A large majority of investigations into RH have dealt with the *real* component of the zeta function but for the purposes of establishing the nature of the analytic properties of the whole function $\log \xi(s)$, consideration of just the imaginary components, particularly in context of the argument as dealt with in this paper is sufficient.

Comparing results of the closed contour with an equivalent *half* contour through symmetry considerations alone reveals a new identity for the argument of (1) – i.e. for $f(s) = \xi(s)$

$$\begin{aligned} \arg f\left(\tfrac{1}{2}+iT\right) &= \arg\sqrt{f\left(-\varepsilon+iT\right)f\left(1+\varepsilon+iT\right)} \\ &= \tfrac{1}{2}\arg f\left(-\varepsilon+iT\right)+\tfrac{1}{2}\arg f\left(1+\varepsilon+iT\right). \end{aligned} \tag{2}$$

This is applicable, not just to any analytic function invariant or symmetric in the variable transformation $s \to 1-s$, but valid for each component in (1) as well, which need not be symmetric, provided one is in the asymptotic regime far up the critical strip, $T >> |\varepsilon|$. To illustrate, $f(s) = \xi(s),\ s(s-1)$ are satisfied by (2) exactly as both are invariant with respect to the variable transformation $s \to 1-s$; however, (2) also applies asymptotically if one were to put $f(s) = \Omega(s),\ \varsigma(s)$ both of which are not invariant under the same symmetry transformation, i.e. $f(s) \neq f(1-s)$:

$$\begin{aligned} \operatorname{Im}\log\Omega\left(\tfrac{1}{2}+iT\right) = {} & \tfrac{1}{2}\operatorname{Im}\log\Omega\left(-\varepsilon+iT\right)+\tfrac{1}{2}\operatorname{Im}\log\Omega\left(1+\varepsilon+iT\right) \\ & +O\left(T^{-1}\right) \end{aligned} \tag{2a}$$

$$\begin{aligned} \operatorname{Im}\log\varsigma\left(\tfrac{1}{2}+iT\right) = {} & \tfrac{1}{2}\operatorname{Im}\log\varsigma\left(-\varepsilon+iT\right)+\tfrac{1}{2}\operatorname{Im}\log\varsigma\left(1+\varepsilon+iT\right) \\ & +O\left(T^{-1}\right). \end{aligned} \tag{2b}$$

For a validation of (2a-b) using a consistent treatment of the argument or phase of a function in terms of the inverse tangent function, see Appendix I and II.

The invariance of (2) to the contour parameter ε is a direct manifestation of the (anti) symmetry inherent in the argument, $\operatorname{Im}\log f(s)$ across the critical strip. Provided the contour is never on a zero, the variation through the critical line is necessarily and sufficiently smooth and continuous. Consequently, for real *T*, and a result of the fact that $\xi\left(\tfrac{1}{2}+iT\right)$ is a real number, the *phase* (argument) $\operatorname{Im}\log\xi(s)$ must then be an integral multiple of π and is therefore a *fixed point* for all *T* between two poles on the critical line. This symmetry allows one to cross a "branch cut" through the critical line without ambiguity as long as the number of zeros up to a height *T* is correctly accounted for, irrespective of the value of the parameter ε.

It is shown that the partner to the argument of the zeta-function is the *winding* function, $\operatorname{Im}\log\Omega^{-1}(s)$. It is smooth, has no zeros or singularities of its own other than to cancel trivial zeros of the zeta function on the negative real line. However it is remarkable that this function effectively counts the zeros of $\varsigma(s)$ on the critical line in a smooth manner, gauging the number around some *expectation* through smoothly *winding* the argument – this can be pictured in a similar manner to Gauss's clock calculator for modular arithmetic of complex numbers instead of integers. In most cases for zeros found lower on the critical strip, it is the nearest integer number to the *exact* number $N(T)$. In contrast, and in concert with that, the argument of the zeta-function, $\operatorname{Im}\log\varsigma(s)$ is fluctuating, or vibrating around this mean or average, $\langle N(T)\rangle$. It cancels the corrections to the exact number of zeros and can be thought of as the error to the actual number of zeros on the critical line, viz. $S(T) := N(T) - \langle N(T)\rangle = \pi^{-1}\operatorname{Im}\log\varsigma(\tfrac{1}{2}+iT)$.

Equation (2b) follows from the fact that in the asymptotic limit (2a) is independent of ε. It is then shown that RH must be true if (2b) is also independent of (or invariant with respect to) the parameter ε, otherwise the value of $\arg\xi(\tfrac{1}{2}+iT)$ would depend on the path taken in the 1st quadrant – either around the off critical zero(s), $\varepsilon > \alpha - \tfrac{1}{2}$ where it picks up a phase, or through the zeros, $\varepsilon < \alpha - \tfrac{1}{2}$ where it does not – see Fig. 1. From this one is forced to conclude that if RH is not true then at some point on the complex plane away from a zero (pole) the phase $\arg\xi(s) = \operatorname{Im}\log\xi(s)$ and hence $\arg\varsigma(s)$ are no longer single-valued and necessarily discontinuous. This is a stronger requirement than what Riemann stated [1], which was that $\varsigma(s)$ must be single-valued; however, it does state precisely consequences of assuming a violation of RH – that is, the argument of the function or, equivalently its square root is non-analytic.

The remedy for ridding oneself of this multi-valued inconsistency and discontinuity in the phase is a branch cut located symmetrically across the critical strip (Figure 1) in the same manner applied to radical complex functions such as $f(s) := \sqrt{s(s-1)}$ so as to make the phase $\arg f(s)$ single-valued between 2 off crucial zeros at 0 and 1 – see also Appendix II. This means that the holomorphic function $\xi(s)$ must also have branch cuts, at points where off critical zeros are to make the phase $\arg\xi(s)$ single-valued, thus contradicting its holomorphic nature. Assuming, for argument one does find a zero(s) off the critical line (in conjugate pairs), and no branch cut is considered necessary, the contour parameter can always be chosen in such a way that it resides between the first ever discovered (hypothetical) off-critical zero(s) uncovered at some distance α from the critical line, viz. $\varepsilon < \alpha - \tfrac{1}{2}$. The implication by the theory of Bohr, Landau [5] is that the zero (or pole in the case of contour integral over diff logs) must come from an external source – in other words be *exogenously* imposed on the original function $\xi(s)$, and therefore not in any way associated with, or affecting the phase of zeroes (poles) found *endogenously* on the critical line.

III. Background, Review and Context

In his paper, Riemann [1] applied a contour integral representation for $\Gamma(s)$ to similarly represent the zeta function in a manner that is valid for all s in the complex plane, except for the solitary point on the real axis, $s=1$

$$\varsigma(s)=\frac{\Gamma(1-s)}{2\pi i}\int_{\infty+}^{\infty-}\frac{(-x)^{s}}{e^{x}-1}\frac{dx}{x} \tag{3}$$

The contour goes counter-clockwise around the origin [1] from $\infty+i0_{+}$ to $\infty-i0_{+}$ in such a way that the branch cut from origin to +∞ is never crossed in order that $(-x)^{s}=e^{s\cdot\log(-x)}$ be single-valued. In this manner the zeta function is uniquely defined and convergent for $s<0$ as well, and he was then able to derive a functional equation for the zeta-function that applies throughout the complex plane, which then can be expressed in a symmetric and holomorphic way through (1). Pitkänen [57] used (3) to find a specific hermitan operator and gauge whereby the proportionally equivalent inner product vanishes implying RH true.

Using another (Mellin) transformation for $\Gamma(s/2)$ instead of $\Gamma(s)$ as above, and with use of Jacobi's identity for the theta function $\theta(x):=\sum_{n=-\infty}^{\infty}e^{-\pi\cdot xn^{2}}=x^{-1/2}\cdot\theta(1/x)$, he proved more rigorously this holomorphic symmetry of the zeta function and went on to derive another alternate representation [1]:[10]

$$\frac{\varsigma(s)}{\Omega(s)}=\Phi(s,\lambda)+\Phi(1-s,\lambda^{-1}) \tag{4}$$

where[11]

$$\Phi(s,\lambda):=-\frac{\lambda^{s/2}}{s}+\sum_{n=1}^{\infty}\frac{\Gamma(s/2,\pi n^{2}\lambda)}{(\pi n^{2})^{s/2}}. \tag{4a}$$

Note the free parameter λ can vary in the range, $0<\lambda<\infty$ while (4) is independent of λ and invariant under the transformation $s\to 1-s$. Riemann showed, by splitting the regimes $0<\lambda\le 1$ and $1<\lambda<\infty$, using the integral representation of the incomplete Γ-function $\Gamma(\mu,x)=\int_{x}^{\infty}e^{-t}t^{\mu-1}dt$ that $\xi(s)$ is *even* in the complex variable $s-1/2$ where $s=\sigma+i\cdot t\ \forall\ \sigma, t\in\Re$ and converges faster than $|s-1/2|^{2}$. The explicit symmetry evident in (4), and the fact that it is invariant with respect to the parameter λ, reveals another important observation that, in the vein of analytic continuation, this kind of functional invariant symmetry may be more prevalent than originally thought – viz. valid for the

[10] There are numerous of ways to prove the functional formula inherent in (4) – see Titchmarsh [14].

[11] Riemann's method can be been generalised with the added a free extra parameter $0<\lambda<\infty$ using the Poisson Summation Formula – see [63] – who Riemann attributed to Jacobi through his theta function identity.

function $\xi(s)$ and its argument $\operatorname{Im}\log\xi(s)$, or equivalently $\sqrt{\xi(s)}$. Castro, Mahecha et al. [59] used (4) to find an operator and gauge where the proportionally equivalent inner product vanishes implying RH true, in precisely the manner as Pitkänen above [57] – for a review of their work see Elizalde et al. [60].

Apparently, Riemann's derivation of (4) was to help prove that (1) is an entire analytic function, which can then be represented by a polynomial of (infinite) degree and thus be expressible as a product over all the non-trivial zeros of the zeta function – viz.

$$\xi(s) = \xi(0)\cdot\prod_{\rho}\left(1 - s/\rho\right) \tag{5}$$

where $\xi(0) = -\varsigma(0) = 1/2$. The product here goes over all the zeros ρ that are known to exist in the (potentially infinite) critical strip $[iT,\ 1+iT]$ $\forall$ T $\in \Re$. Hadamard [15], in 1893 proved the product (5) and more generally using analytic properties of even entire functions and uniform convergence of the sums and products – for detailed discussions see Edwards [16]. As (5) has been proven and known to be absolutely convergent, one can then take the natural logarithm of (1) and differentiate term by term to obtain

$$\frac{\xi'(s)}{\xi(s)} = \sum_{\rho}\frac{1}{s-\rho} = \frac{1}{s-1} + \frac{1}{s} + \tfrac{1}{2}\psi\left(\tfrac{s}{2}\right) - \tfrac{1}{2}\log\pi + \frac{\varsigma'(s)}{\varsigma(s)} \tag{6}$$

Here, $\psi(s)$ is the digamma function. From (5) and (6) it is understood that the zeros are ipso facto expressible as simple poles including any potential multiplicities.[12] While the zeros are explicit in the middle term they are implicit in the right hand side of (6) – all zeros, including any potential (off critical) ones that violate RH must be *endogenous* and generated from the zeta function $\varsigma(s)$. For an in-depth description of the analytic properties of a function representing *exogenous* off critical zeros in the complex plane, see Appendix II.

Applying Cauchy's Integral Theorem to Eq.(6) and using the complex contour as shown in Figure 1, the number of zeros in a specific vertical range of the upper critical strip in the range $0 < \operatorname{Im}(s) \le T,\ 0 < \operatorname{Re}(s) < 1$ is precisely determined by the integral

$$N(T) = \frac{1}{2\pi i}\oint_{c_0+c_1+c_2+c_3} ds\,\frac{\xi'(s)}{\xi(s)}. \tag{7}$$

The closed contour comprises four connected broken line segments controlled by an adjustable parameter $\varepsilon > -1/2$ measuring width of the closed contour: $c_0 = [-\varepsilon, 1+\varepsilon]$, $c_1 = [1+\varepsilon, 1+\varepsilon+iT]$, $c_2 = [1+\varepsilon+iT, -\varepsilon+iT]$ and $c_3 = [-\varepsilon+iT, -\varepsilon]$. Now, the integral (7) can be split into its separate components and analysed individually. Provided no zeros (or poles in this case) reside on the contour, (7) captures any and all zeros located within an

[12] Historically, researchers have been compelled to explicitly verify that all non-trivial zeros up to a certain level T are simple

arbitrary region such as the critical strip, which is known to be where all of them exist. For starters contribution to (7) from the integration contour c_0 is identically zero by symmetry.

Consider next the integral component across the critical strip at constant height T

$$\int_{c_2} d\log\xi(s) = \log\xi(-\varepsilon + iT) - \log\xi(1+\varepsilon+iT) \tag{8}$$

From the functional relation (1) the real component of $\log\xi(s)$ must be even with respect to the transformation $s \to 1-\bar{s}$ where $\bar{s}$ is the complex conjugate of $s = \sigma + iT$ and only the *real* component of *s* varies. By symmetry variation along the contour c_2 in the integral (8) must be purely imaginary – i.e. $\mathrm{Re}\int_{c_2} d\log\xi(s) = 0$.

Using the standard asymptotic formula for large modulus $|s| \sim T >> 1$ of the logarithm of the Γ- function

$$\begin{aligned}\log\left[\Omega^{-1}(s)\right] &= -\log\Omega(s) \\ &= \left(\frac{s-1}{2}\right)\log(s) - \frac{s}{2}\log(2\pi e) + \tfrac{1}{2}\log\pi + O\left(T^{-1}\right)\end{aligned} \tag{9}$$

in (1) and (8), and keeping in mind the even symmetry of the real component of $\log\xi(s)$, one is left with

$$\begin{aligned}&\mathrm{Re}\log\varsigma(-\varepsilon+iT) - \mathrm{Re}\log\varsigma(1+\varepsilon+iT) \\ &= \left(\frac{1}{2}+\varepsilon\right)\log\left(\frac{T}{2\pi}\right) + O\left(T^{-2}\right),\end{aligned} \tag{10}$$

which is a result of the Lindelöf's Theorem concerning the growth of the modulus of the zeta-function, and should be compared with Eq.(1) sect. 9.2 of Edwards [16]:

$$\left|\frac{\varsigma(-\varepsilon+iT)}{\varsigma(1+\varepsilon+iT)}\right| \sim \left(\frac{T}{2\pi}\right)^{\frac{1}{2}+\varepsilon}. \tag{10a}$$

In the same way the real component is even, the imaginary component or argument[13] $\arg\xi(s) = \mathrm{Im}\log\xi(s)$ must be odd with respect to the variable transformation above,

[13] This definition of argument is more general than *arc tan* as its value is not necessarily bounded, provided it is smooth or (piecewise) continuous – see Appendix I

$s \to 1-\bar{s}$. Again, using the asymptotic formula (9) in (8) gives for the imaginary part of the zeta function, a result that is not quite so definitive as (10) – viz.

$$\begin{aligned} &\arg\varsigma(-\varepsilon+iT)-\arg\varsigma(1+\varepsilon+iT) \\ &\quad=\tfrac{\pi}{4}(1+2\varepsilon)-i\cdot\int_{c_2} d\log\xi(s)+O\left(T^{-1}\right) \end{aligned} \tag{11}$$

As described previously, the real part of the integral on the right hand side of (11) is *zero* and therefore $-i\cdot\int_{c_2} d\log\xi(s)=\arg\xi(-\varepsilon+iT)-\arg\xi(1+\varepsilon+iT)$. It is not known what if any bounds there are on the modulus of (11) but is known to oscillate and grow very slowly as T increases. It is also directly tied to the truth of RH or otherwise – see Appendix I.[14]

Next, consider the integral through the real axis and adjacent to the critical strip

$$\begin{aligned} \int_{c_1'=c_1+c_3} d\log\xi(s) &= \log\xi(1+\varepsilon+iT)-\log\xi(1+\varepsilon-iT) \\ &= i\cdot 2\arg\xi(1+\varepsilon+iT), \end{aligned} \tag{12}$$

where the contour c_3 is redefined according to the variable transformation $s \to 1-s =$ $1+\varepsilon+iT$ to make a new contour integral $c_1'=[1+\varepsilon-iT, 1+\varepsilon+iT]$ that is equivalent to c_1+c_3 as one long vertical line through the real axis *along*side ($\varepsilon \geq 0$), and perhaps *through*out ($-\frac{1}{2}<\varepsilon<0$), the critical strip – see Fig. 1. In contrast to (8), where the real component changes *across* the critical strip, variation in (12) involves only the imaginary component of s. However, the same (anti) symmetric property evident in (8) is also a property that is shared by (12) – implying both are purely imaginary. Combining the integrals (8) and (12) and, keeping in mind its equivalence to (7), reveals that any and all (simple pole) singularities including any multiplicities representing the non-trivial zeros up to a height T are contained in the critical strip, and provided that no zeros reside on the integration contour defined by ε.[15]

Combining (11) with a similar substitution of (9) into (12) gives

$$\begin{aligned} \int_{c_1'} d\log\xi(s) = i\cdot\left(T\log\left(\frac{T}{2\pi e}\right)+2\pi\left(1+\frac{\varepsilon}{4}\right)+2\arg\varsigma(1+\varepsilon+iT)\right) \\ +O\left(T^{-1}\right), \end{aligned} \tag{13}$$

---

[14] For details of pioneering work up towards the middle and latter part of the last century, see Edwards [16] and Titchmarsh [14].

[15] As mentioned the controlled parameter, ε which defines the horizontal position of the vertical integration contour not on a zero, is adjustable from $-1/2$ to $+\infty$ – see Figure 1.

therefore, (11) combined with (13) is the same as (7):

$$N(T) = \frac{T}{2\pi}\log\left(\frac{T}{2\pi e}\right) + \frac{7}{8} + O\left(T^{-1}\right)$$
$$+\frac{1}{\pi}\arg\sqrt{\varsigma\left(-\varepsilon+iT\right)\varsigma\left(1+\varepsilon+iT\right)} \tag{14}$$

This is the standard result [1,17] with one important caveat that, at least asymptotically for $T >> \varepsilon$ to $O\left(T^{-1}\right)$, the following identity holds:

$$\arg\varsigma\left(\tfrac{1}{2}+iT\right) \approx \arg\sqrt{\varsigma\left(-\varepsilon+iT\right)\varsigma\left(1+\varepsilon+iT\right)}$$
$$= \tfrac{1}{2}\arg\varsigma\left(-\varepsilon+iT\right) + \tfrac{1}{2}\arg\varsigma\left(1+\varepsilon+iT\right).$$

This identity shows an invariant symmetry relationship. It is observed in many forms – in one example a similar symmetry relation involving a different functional representation and mapping of parameters $s \to 1-s$ vs. $\lambda \to \lambda^{-1}$ is found in (4). Another example is found in identity (A7) of Appendix I. Of direct relevance to the above, and showing precisely same symmetry, is the inner *scalar* product proportional to either $\varsigma(s)$ [57] or $\varsigma(s)\cdot\Omega^{-1}(s)$ [59] associated with conjugate states or Hilbert-Polya operator that are precisely the orthogonal states manifest as zeta zeroes on the line, implying RH is true – compare (4,4a) with Eqs. (6), (11)-(17) of ref. [60]. This is analogous to the generally non-orthogonal states of the Berry phase or Wess-Zumino term in WZW theory, which is just the NLσM associated with (invariant) super symmetric conformal theory as touched on in the Introduction.

In the next section a rigorous formulation of this non-*trivial* symmetry is demonstrated, particularly in context of the *ξ*-function (1) and to its component factors and other analytic functions in general. It is proposed that, considering the (piecewise) continuous, single-valued *analytic* nature of the function(s) in question, this (asymptotic) symmetry provides the basic, fundamental *universal* criteria required in establishing the truth of the Riemann Hypothesis and consequences to the contrary.

IV. Formulation

By virtue of symmetry alone, one need take only half the contour above and below the real axis c_1', viz. c_1 and half the contour c_2 on either side of the critical line in (7), viz. $c_2' = \left[\frac{1}{2}+iT, 1+\varepsilon+iT\right]$ and consider only the imaginary component of the resulting integral.[16] Assuming no zeros reside *on* the contour this integral contributes precisely half the full integral and therefore counts exactly half the number of zeros enclosed within the full adjustable contour, regardless of whether any off-critical, yet unaccounted for, zeros are enclosed or otherwise. Using these facts one can recast (7) slightly differently

$$N(T) = \frac{1}{2\pi i} \int_{c_1'+c_2} d\log\xi(s) = \frac{1}{\pi} \operatorname{Im} \int_{c_1+c_2'} d\log\xi(s) \tag{15}$$

By the fundamental theorem of calculus the half contour, as represented by the right hand side of (15), can be written as

$$N(T) = \frac{1}{\pi}\arg\xi\left(\tfrac{1}{2}+iT\right) = 1 + \frac{1}{\pi}\left[\vartheta(T) + \arg\varsigma\left(\tfrac{1}{2}+iT\right)\right]. \tag{16}$$

Here, $\vartheta(T) := \arg\Omega^{-1}\left(\frac{1}{2}+iT\right) = -\arg\Omega\left(\frac{1}{2}+iT\right)$ is recognised as the Riemann-Siegel Theta function. By virtue of the fact that $\xi\left(\frac{1}{2}+iT\right)$ is real, it is implied in (16) that $\pi^{-1}\arg\xi\left(\frac{1}{2}+iT\right)$, within the (half) contour $c_1 + c_2'$, is a real positive integer and does not pick up any *explicit* dependence on the integration parameter ε. In contrast, an apparent dependence on ε does appear for the full contour integral considered in the middle term of (15), representing the full contour

$$N(T) = \frac{1}{2\pi}\arg\xi(-\varepsilon+iT) + \frac{1}{2\pi}\arg\xi(1+\varepsilon+iT). \tag{17}$$

One expects, that by symmetry this ε - dependence should cancel out. Using the results in the previous section by comparing (14) with (16), we find that for large *T*,

$$\begin{aligned} &\tfrac{1}{2}\arg\Omega^{-1}(-\varepsilon+iT) + \tfrac{1}{2}\arg\Omega^{-1}(1+\varepsilon+iT) \\ &\quad = \vartheta(T) + O\left(T^{-1}\right) \end{aligned} \tag{18}$$

is single-valued and independent of the parameter ε in the asymptotic limit. Next, one can apply (16), (17) and (18) into (15) for each term in the logarithm of (1). This then implies

[16] In 1912, Backlund [17] considered this (imaginary component of the) half contour integral, with contour parameter $\varepsilon = \frac{1}{2}$ – see also Edwards [16]

$$\arg\varsigma\left(\tfrac{1}{2}+iT\right)=\tfrac{1}{2}\arg\varsigma\left(-\varepsilon+iT\right)+\tfrac{1}{2}\arg\varsigma\left(1+\varepsilon+iT\right) \\ +O\left(T^{-1}\right). \tag{19}$$

In fact, the following is exactly true for $f(s)=\xi(s),\ s(s-1),\ \varsigma(s)\cdot\Omega^{-1}(s)$ by virtue of the ε - invariant (even) symmetry obeyed by $f(s)=f(1-s)$:

$$\arg f\left(\tfrac{1}{2}+iT\right)=\arg\sqrt{f\left(\tfrac{1}{2}+\delta+iT\right)\cdot f\left(\tfrac{1}{2}-\delta+iT\right)} \\ =\tfrac{1}{2}\arg f\left(\tfrac{1}{2}+\delta+iT\right)+\tfrac{1}{2}\arg f\left(\tfrac{1}{2}-\delta+iT\right) \tag{20}$$

For convenience we have redefined the contour path parameter $\delta:=\tfrac{1}{2}+\varepsilon$, valid for any $0<\delta<\infty$ in such a way that no zeros reside on the contour. Note that (20) is valid to $O\left(T^{-1}\right)$ for $T>>\delta$ when $f=\Omega,\ \varsigma,\ s, s+\delta, \delta s$ ... etc. as well, as is evident in (18) and (19), even though $f(s)\neq f(1-s)$. A closer examination of the argument / phase as represented by the inverse tangent function for variables originally defined in the 1st quadrant and dynamically continued to the other 3 quadrants can be found in Appendix I. One particular case where (20) may be violated, comprising a broken symmetry of sorts for the simple case $f(s)=\left(s-\rho_\alpha\right)\left(s+\rho_\alpha^*-1\right)$, is illustrated in Appendix II. Through a certain judicious adjustment of the contour parameter δ for a fixed *T*, this can always be treated consistently to a single isolated violation of RH excluding multiplicities.

If any off-critical zeros exist they must arise *endogenously* – i.e. generated within the fluctuation term (19) and cannot contribute to, or be counted by, the smoothly varying non-fluctuating component $\vartheta(T)$, which as prescribed above gives a measure or "expectation" of number of zeros on the critical line only. This means that (19) is either multi-valued or untrue. Historically, this fluctuation term has been likened to the definition $S(T):=$ $N(T)-\pi^{-1}\vartheta(T)-1$, which when combined with (16) is the same as $\pi^{-1}\arg\varsigma\left(\tfrac{1}{2}+iT\right)$. The location of the horizontal position of the vertical contour, as measured by the parameter, ε as described here is adjustable and serves to illuminate any inconsistencies in the analytic nature of the function ξ(*s*) by isolating a first ever discovered *hypothetical* off critical zero when assuming RH is false.

One is now in a position to see what consequences there are in considering whether RH is true or not. For a number of reasons, focus has been made primarily on the *real* rather than on the *imaginary* component of the log of the zeta function as a result of the fact that both the components through the modulus are accounted for in locating zeros with the *real* logarithm of the zeta function. On the other hand, the *imaginary* component is easier to deal with, in spite of pitfalls associated with the multi-valued nature of *argument* or *phase* – examples include Schlömilch [64] series of sums that, in one representation are quite complex but manifest as generalisations of simple piecewise continuous functions associated with Fourier series when considering only the imaginary components of lattice sums. From the analysis described here, based purely on symmetry, one must conclude that if RH true then $N(T)$, as defined by either full or half contour in (15), will have no dependence on the parameter ε in any way – i.e. is *universally* invariant.

V. Consequences of RH or otherwise

Conditions under which (16)-(19) are independent of the integration parameter ε or otherwise are investigated in detail in what follows. In this context there's no need to consider the distribution of zeros other than to note that $\vartheta(T) := \arg \Omega^{-1}\left(\tfrac{1}{2}+iT\right)$ gives some non-fluctuating "expectation" of the *gauge* or winding the number of zeros residing on the critical line $\mathrm{Re}(s)=\tfrac{1}{2}$ up to some height T on the critical strip; as such, $\vartheta(T)$ by invariance (or symmetry) doesn't register or accumulate any potential off-critical zeros – those zeros must then be captured by $S(T)=\pi^{-1}\arg\varsigma\left(\tfrac{1}{2}+iT\right)$, which *fluctuates* in value or vibrates around some mean value.

The nature of the distribution of zeros on the critical line or otherwise is a completely different matter as to whether they reside on the critical line, and relates to the fact that Gram points g_n as defined by $\vartheta(g_n) := \pi n$ can give good proxies for locating actual zeros described through the error term $S(T)=N(T)-\langle N(T)\rangle$, whereby[17] $\langle S(T)\rangle = 0$, and therefore

$$\langle N(T)\rangle = \frac{\vartheta(T)}{\pi}+1 . \tag{21}$$

Gram's so-called "law" effectively states that the actual number of zeros on the line is precisely the nearest integer given in (21) insofar as there is always a zero on the line between two Gram points g_n and g_{n+1}. Of course, exceptions to this rule do occur if rarely, yet with increasing frequency the higher up the critical line one goes. This is essentially equivalent to proving that $\langle \ldots \rangle =$ "expectation"; for example,

$$\langle S(T)\rangle = \lim_{T\to\infty}\frac{1}{T}\int_0^T S(T)\,dt := \langle N(T)\rangle - \frac{\vartheta(T)}{\pi} - 1 = 0 \tag{22}$$

or

$$\langle S(T)\rangle = \lim_{T\to\infty}\frac{1}{N(T)}\sum_{n=1}^{N(T)}\frac{1}{\pi}\,\mathrm{Im}\log\varsigma\left(\tfrac{1}{2}+it_n\right) = 0 , \tag{23}$$

[17] A result proved by Littlewood [3]. Turing [8] proved $\forall\, n \in \{0,1,2\ldots\}, S(g_n)=0$ and related work that greatly enhanced the finding of zeros on the critical line – for discussions see Edwards [16].

Note that, even if the distribution is not bounded the expectation is well defined. Here, $N(T) = \sum_{n=1}^{N(T)} = \sum_{n=1}^{\infty} \Theta(T - t_n)$ where $\Theta(x)$ is the Heaviside Step function and t_n is the height of the n^{th} zero on the line.

As mentioned the nature of the distribution of zeroes on the line or otherwise is completely separate from the following consideration that shall demonstrate consequences of RH being true or false, up to some level or height up the critical strip $\operatorname{Im} s = T$. This is due to the fact that any contour cannot strictly speaking cross the critical line, which itself if a branch cut, however due to the fact that $\xi(\frac{1}{2} + iT)$ is real, its argument is always a fixed point $i\pi \cdot n$ (unchanging) for all *T* between two zeros on the line. A key point to consider is that the zeros must be generated from some source, either internally from $S(T)$ or externally (*exogenously*) to the function,[18] very much as would be from adding the simple function $f(s) = s(s-1)$, but instead through adding to (1) some another function comprising two extra zeros far up and within the critical strip. This is explicitly manifest as

$$f(s) = (s - \rho_\alpha)(s + \rho_\alpha^* - 1), \tag{24}$$

where

$$\rho_\alpha = \tfrac{1}{2} + \alpha + it_\alpha \tag{24a}$$

is the very first hypothetical off-critical zero found on the critical strip a distance α to the right of critical line up to some height t_α, while ρ_α^* is merely the complex conjugate symmetrically represented below the real axis but not covered by the contour – for a detailed treatment and discussions around consequences to RH, see Appendix II.

We consider two cases separately:

i) RH true up to some specified height on the critical strip: $T < t_\alpha$

ii) RH false. First hypothetical off-critical zero captured: $T > t_\alpha$

In case i) there are no big surprises except that because $\xi(s)$, as defined in (1), is holomorphic and the complex contour integration parameter ε can vary within the range $-\frac{1}{2} < \varepsilon < \infty$ without changing or affecting the result. Therefore, $N(T)$ is just the number of zeros on the critical line and (16) is single valued, which is the same as saying that (17) is independent of ε. This is essentially the situation we find ourselves in today, up to a point of some vast number of zeros, $\sim 10^{22}$ by current accounts – see Odlyzko [19].

For Case ii) the half contour encloses the 1st hypothetically discovered (pair of) zeros(s) generated by the fluctuation term $S(T)$ that measures the error estimate between the expected and actual number zeros on the line up to some height *T*. By Cauchy's Integral

---

[18] An internal manifestation of zeros is termed *endogenous* while an external imposition is *exogenous*

theorem, (19) depends on the whether the contour parameter $\varepsilon = \delta - \frac{1}{2}$ captures the zero $(\delta > \alpha)$ or does not $(\delta < \alpha)$ – i.e. the number of zeros captured is

$$N(T)\big|_{\delta>\alpha} = N(T)\big|_{\delta<\alpha} + m \tag{25}$$

where $m \in \{0,1,2, \ldots\}$ is the *multiplicity* of the zero (or *residue* of the pole) contained in the half (or full) contour $c_1 + c_2'$ ($c_1' + c_2$, respectively). As the uniform "expectation" or mean derived from full contour $\left\langle N(T)\big|_{\delta>\alpha} \right\rangle = \left\langle N(T)\big|_{\delta<\alpha} \right\rangle$, as represented by the average winding of the phase (21), is not δ-dependent – i.e. it doesn't account for any off-critical zeros – then the only contribution must come from the phase of the zeta function itself

$$S(T)\big|_{\delta>\alpha} = S(T)\big|_{\delta<\alpha} + m\,. \tag{26}$$

Comparing (26) with (16) means that, if the zero is *endogenous* it must come from the fluctuating term $S(T)$, which is no longer single-valued. This is the same as saying that (19) depends on ε to leading order in the form of an implied Heaviside Step function $m \cdot \Theta(\delta - \alpha)$. In order to preserve the single-valued nature of $\varsigma(s)$ [1], one must have an *even multiplicity* of the zero enclosed within the half contour at all times; that is to say, m $\in$ {0, 2, 4, …} in (25) and (26). Since the position of the vertical contour can be adjustable anywhere in the critical strip up to some height *T*, then a single zero (pole) can always be isolated; therefore, at least a double multiplicity is required to capture a single zero and for $\varsigma(\frac{1}{2} + iT)$ to be single-valued. This also ensures that the continuous analytic and invariant symmetry of (20) is maintained for the function $f(s)$, its argument and its square root –see Appendix II, which gives an example of a function where this cannot be the case unless a branch cut is maintained, implying that the original function wasn't holomorphic or analytic to begin with.

This is not unrelated to Turing's result [8] that $S(g_n)$ must be an even integer or that $e^{i\vartheta(T)} \cdot e^{i\pi S(T)} = \pm 1/\pm 1 \,\forall\, T \in \Re$, even though one is not just talking about zeros on the critical line but throughout the complex plane, as is plainly evident from the symmetrical mapping relationship between the full and half contour integral (15), which then establishes the symmetry relation (20) for the argument of $f(s) = \xi(s)$ or $\varsigma(s)$. Having a single off-critical zero with at least multiplicity $m = 2$ in (26) makes it severely less probable that RH is false as the change in value of $\arg\varsigma(s)$ across the critical strip $[1+iT,\ iT]$ is typically small – for more discussions, see Appendices I and II. By measuring the change in phase / argument across the critical strip where the zero is *on* the contour, as in (23), and equating the number to a half-odd integer (between the contour enclosing and including zero), as described in (11), one sees that the error term $S(T)$ grows very slowly as a function of T. As an initial check the data provided by Odlyzko [19] for the 10,000 zeros from 10^{12} show that in none of these cases does the error term have modulus greater that 1.5 – see Appendix II.

Notice that, if one were to impose zeroes is onto the function $\xi(s)$, *exogenously* as it were, then there would be no inconsistency in the symmetry or the single-valued (piecewise) continuous nature of the phase/argument of the Riemann Zeta function as long as the zeroes reside on the critical line $\operatorname{Im}\log\varsigma\left(\frac{1}{2}+iT\right)$. By imposing a stronger condition than Riemann's that *both* $\xi(s)$ and $\arg\xi(s)$, or more specifically that $\varsigma(s)$ and $\arg\varsigma(s)$ are analytic, single-valued &/or piecewise continuous then RH must be true. In other words if RH true then a holomorphic function and its natural logarithm (apart from points were zeros reside) can both be analytic functions in the entire complex plane – i.e. s is defined unambiguously within all 4 quadrants of the Cartesian plane. If RH false then $\arg\sqrt{\xi(s)}$ is necessarily discontinuous due to path dependence as represented by the contour integral, either by capturing $\delta > \alpha$ or not capturing $\delta < \alpha$ the zero, requiring the presence of branch cuts.

Appendix I – The Argument and Inverse Tangent Function

Preposterous Proposition: There exist conditions under which $-1/1 \neq 1/-1$

Any good student in intermediate school would consider the statement above to be untrue simply by multiplying both sides by –1 and conclude that $(-1)^2 = 1$. Well done, but what happens if one decides to taken the square root of the above non-equality? Well, the left hand side would equal *i*, which by definition is just the square root of minus one, while the right hand side would equal to $1/i = -i$, and so therefore the non-equality holds. So which is true? By the simplest example just described, it is the one that includes the possibility of both, depending on the context. Traditional complex analysis deals with this apparent inconsistency by applying a branch or cut line in the complex (or conformal) plane, where integration contours are forbidden to cross so that the analytic function in question remains single-valued and (piecewise) continuous. By allowing for the preposterous above, a need for imposing branch cut structures *ex post* to the logarithm or *argument* function considered here is revealed to be unnecessary, provided the context of the specific initial conditions and boundaries is properly accounted for.

The purpose here is to put into context a consistent treatment of the *argument* of an arbitrary complex analytic function, $f(s)$ of a complex variable $s = x + i\,y$ in terms of the inverse tangent function. The basic definition is

$$\arg f(s) := \operatorname{Im}\log f(s). \tag{A1}$$

For any analytic function expressed in polar coordinates $f(s) = R(s)e^{i\theta(s)}$, (A1) can be written as the inverse tangent:

$$\arg f(s) = \theta(s) = \arctan\frac{\operatorname{Im} f(s)}{\operatorname{Re} f(s)}, \tag{A2}$$

Obviously the analytic properties of (A1, A2) depend on the nature of the function in question and can potentially be arbitrarily defined or *multi-valued* if one does not take the necessary care. The context may be arbitrary as to whether it corresponds to the principal value, implying a range of applicability in the 1st and 4th quadrant only – i.e. $\operatorname{Re} f(s) > 0$

$$-\pi/2 < \arctan(t) < \pi/2, \qquad -\infty < t < \infty, \tag{A3}$$

or whether it applies to other ranges such as the 2nd and 3rd quadrants of the complex plane, $\operatorname{Re} f(s) < 0$. The exercise here is merely to illustrate a very simple example of how one can treat (A2) uniquely as well as (piecewise) continuously in the complex plane and show what also may occur for the more involved function $\xi(s)$ treated in the main section. Details of the elementary case $f(s) = s$ can be gotten from G&R 1.6 [65].

Consider two angles defined on a triangle in the 1st quadrant of Descartes *x-y* plane as shown in Figure 2 and apply the same elementary technique used to uniquely define the standard trig functions for arguments in all 4 quadrants but instead apply it to the inverse tangent function so that it too may be defined for the 2nd and 3rd quadrants as well.

Looking at the trivial case of the angles defined in the 1st quadrant – see Figure 2

$$\theta(s) = \arctan\frac{\operatorname{Im}(s)}{\operatorname{Re}(s)} = \arctan\left(y/x\right), \quad 0 < \theta < \pi/2 \tag{A4}$$

$$\varphi(s) = \arctan\left(x/y\right), \quad x,\ y > 0, \qquad 0 < \varphi < \pi/2. \tag{A5}$$

one has by definition $\theta + \varphi = \pi/2$. The question is: does this apply when $\operatorname{Re}(s) = x < 0$. In other words, does the following formula hold in all 4 quadrants – see G&R 1.6.32.2 [65]:

$$\arctan\left(y/x\right) + \arctan\left(x/y\right) = \pi/2, \qquad -\infty < x,\ y < \infty. \tag{A7}$$

To prove (A7) let's look at each case, one by one. Now that one trivially established (A7) in the 1st quadrant (where *x and y* > 0), consider the case where $x > 0$ as before but now $y < 0$ – i.e. one's in the 4th quadrant – see Fig. 2. Define new angles $\theta,\ \varphi$ as (symmetric) transformations of those defined in the 1st quadrant, by a simple negative rotation or reflection in real axis $\theta \to -\theta$ implying a positive rotation or reflection in the imaginary axis $\varphi \to \pi - \varphi$, where one finds that $-\pi/2 < \theta < 0$ and $\pi/2 < \varphi < \pi$. In this context, again by looking at Fig. 2, we can write the first term in (A7) as

$$\theta = \arctan\left(-|y|/x\right) = -\theta_0 \tag{A8}$$

and the second term as

$$\varphi = \arctan\left(x/-|y|\right) = \pi - \varphi_0, \tag{A9}$$

which proves, by adding (A8) to (A9), that (A7) is valid for $y < 0$ as well. Here, any symbols with subscript zero correspond to those angles originally defined on the 1st quadrant as in (A4) and (A5). Note that we have established that the inverse tangent function in (A9) can be defined continuously beyond the range provided in (A3).

Next, consider the case where $y > 0$ as in 1st quadrant but now $x < 0$ – i.e. we're in the 2nd quadrant. By symmetry the exact same conditions apply as for the 4th quadrant as in (A8) and (A9) but with $\theta \to \varphi$ and $x \to y$ wherein $\pi/2 < \theta < \pi$ and $-\pi/2 < \varphi < 0$. This extends the range of validity of (A7) to all $-\infty < x < \infty$ *or* $-\infty < y < \infty$ not both – i.e. in the 1st, 2nd and 4th quadrants based on the definition provided and evident in Fig. 2.

In either case, the preposterous proposition is valid – i.e. one cannot always treat a negative sign in the denominator of the inverse tangent function in the same way as the numerator, since for $x, y > 0$

$$\arctan\left({}^{-y}\!/_{x}\right) = -\arctan\left({}^{y}\!/_{x}\right), \tag{A10}$$

$$\arctan\left({}^{y}\!/_{-x}\right) = {}^{\pi}\!/_{2} + \arctan\left({}^{-x}\!/_{y}\right). \tag{A11}$$

Lastly, consider the case where both $x < 0$ *and* $y < 0$. We're now in the 3rd quadrant. One may take a continuous path from the 1st to the 3rd quadrant following via either the 4th or 2nd quadrant. Consider taking the latter (counter clockwise) transformation from the 1st to 2nd and then from 2nd to 3rd quadrants in Fig. 2 – i.e. applying $\theta \to {}^{\pi}\!/_{2} + \theta$ or $x \leftrightarrow y$ twice in (A8, A9). Using (A10, A11) we find the value appropriate to the 3rd quadrant in the positive sense, $\theta = \pi + \theta_0$:

$$\theta = \arctan\left({}^{-|y|}\!/_{-|x|}\right) = {}^{\pi}\!/_{2} + \arctan\left({}^{|x|}\!/_{-|y|}\right) = \pi + \theta_0. \tag{A13}$$

It doesn't matter whether one chooses θ to progress in the positive sense or vice versa since in either case, relative to θ, φ always evolves in the opposite sense. The ambiguity arises from an often made mistake of choosing both progressions at once – i.e. getting to the 3rd quadrant from the 4th and from the 2nd. In order to prove that (A7) is valid for all $-\infty < x, y < \infty$, based on the definition provided one has to make a choice of, say the path of least resistance. Making another choice and progress θ_0 in the negative sense then repeating the same process as above for (A8-11) will cause problems – viz.

$$\begin{aligned}\theta = \arctan\left({}^{-|y|}\!/_{-|x|}\right) &= -{}^{\pi}\!/_{2} + \arctan\left({}^{-|x|}\!/_{|y|}\right). \\ &= -{}^{\pi}\!/_{2} - \varphi_0 = -\pi + \theta_0.\end{aligned} \tag{A14}$$

This of course implies that the inverse tangent is no longer single valued requiring a branch cut. If we're only interested in cases where the absolute change in phase (argument) is less than 2π this will guarantee to cover all 4 quadrants uniquely. So in choosing θ_0 to progress in the positive sense the choice for the inverse tangent function in the 4th quadrant must be (A13); however, it must be then be true that the correct choice for $\varphi = \arctan\left({}^{-|x|}\!/_{-|y|}\right)$ in the 4th quadrant is in the opposite sense – which is to say that (A14) is appropriate for $\varphi \to \theta$ and $x \leftrightarrow y$ if one chooses (A13) for θ. This completes the proof of (A7) for all 4 quadrants in the static or limiting case for the argument of $f(s) = s$ considered here.

The dynamic case is more interesting – for instance, consider (a) $x > y$ in the 1st quadrant, where one observes from Fig. 2 that $\theta_0 < \pi/4$ & $\varphi_0 > \pi/4$. Then apply (A13) to $\theta = \pi + \theta_0$ and (A14) to $\varphi = -\pi + \varphi_0$ implying that $\pi < \theta < 5\pi/4$ and $-3\pi/4 < \varphi < -\pi/2$. Adding the two results gives $\theta + \varphi = \pi/2$, the same invariant relation as before. Taking the difference gives $\Delta := \theta - \varphi < 2\pi$ and everything is continuous and consistent within the range defined above so again nothing new here. If instead we assume (b) $x < y$ and we are still moving in the positive direction then $\theta \to \pi + \theta$ and one has $\theta_0 > \pi/4$ and $\varphi_0 < \pi/4$. This implies that $5\pi/4 < \theta < 3\pi/2$ and $-\pi < \varphi < -3\pi/4$. Again, $\theta + \varphi = \pi/2$ is invariant as represented by (A7) but the change in angles or angular rotation $\Delta := \theta - \varphi > 2\pi$ is actually *outside* the range 2π but is adequately covered within just 4 quadrants!

The fact that there is an overlap of angles in the 3rd quadrant may seem at first glance to indicate another ambiguity in definition or imply a multi-valued argument. However, this is not the case. The invariant relation (A7) as originally defined in the 1st quadrant may apply unambiguously to the 3rd quadrant as well in the context of continuity of the inverse tangent function with regard to continuous x in relation to a continuous y in the entire plane. It's interesting that there is a potential singularity of the argument at the origin that would be captured by some arbitrary undefined phase corresponding to $\mathrm{Re}\, s = \mathrm{Im}\, s = 0$; in the main section this corresponds precisely to the Riemann zeroes: $f(s) = \xi(s)$.

This result, while basic, is important and remarkable for it illustrates that, not only can one remove any need for branch cuts in context of the function in question and get around the multi-valued traps of the argument that may be stumbled onto via (A13, 14) if one's not careful, but one can go beyond the naïve assumption that within the 1st four quadrants in Descartes' x-y plane a change in *dynamical* variable or argument can only be defined within 2π. This constraint is not required in the recipe given above and is based on the non-equality of the *proposition*, sidestepping the need for imposing branch cuts used in the traditional treatment of the *arc tan* and other functions. The invariant relationship between the angles θ and φ, represented in (A7) for all 4 quadrants serves as a direct analogue to equations (18)-(20) for constant T (or energy levels) across the critical strip.

That the dynamical change in angles Δ is allowed beyond range of 2π is clearly evident when considering the function $f(s) = \xi(s)$ instead – i.e. change in values of $\arg \xi(s)$ across the critical strip near two extremely closely situated zeros on the critical line $\mathrm{Re}(s) = 1/2$. What can happen is that near the two zeros the variation (or curvature) of $\arg \xi(s)$ across the critical strip is such that one finds[19]

$$\Delta = \left| \int_{c_2} ds \cdot \frac{\xi'(s)}{\xi(s)} \right| > 2\pi \,. \qquad \text{(A15)}$$

The inequality (A15) has *direct* consequences for RH being difficult to prove for if this were not the case, then RH would necessarily be true. The key is to prove that RH is true as a consequence of having $\arg \xi(s)$ single-valued and analytically smooth or piecewise

[19] This is indicative of Lehmer's Phenomenon – for discussion see Edwards [16].

continuous. Appendix II looks at an *exogenous* representation of what one may expect if RH false.

Appendix II – Exogenously Defined Off-critical Zeros

Consider the symmetric function $f(s) = f(1-s)$:

$$f(s) := (s-\rho_\alpha)(s+\rho_\alpha^*-1)(s-\rho_\alpha^*)(s+\rho_\alpha-1) \qquad \text{(B1)}$$

where $\rho_\alpha = \tfrac{1}{2}+\alpha+i\cdot t_\alpha$ represents one of four symmetrically placed zeros about the real axis and across the line of symmetry $\operatorname{Re}(s)=\tfrac{1}{2}$. These zeroes may represent the same hypothetical 1st ever discovered off critical zero(s) considered in this paper as potential counterexamples to RH as captured through adjusting the contour around the zero or pole. The only difference is that the zeros considered in (B1) are explicit or *exogenous* – i.e. defined externally in the complex plane and imposed by adding to (or subtracting zeros from) both the sum and right hand side of (6). In contrast the off critical zeros considered in the main section must be implicit or *endogenous* – that is, that they are generated internally through $\varsigma(s)$ in the critical strip as represented by the right hand side of (6) and not imposed from the *outside* as it were.[20] In the first case, zeros are added to (subtracted from) the right hand sides of (6) while in the latter case they are not. One hopes to gain insight as to what are the consequences of RH being false by looking at the explicit form as given in (B1), which is covered in detail here.

Again, apply precisely the same contour integral as considered in the main part of this paper. It is clear through symmetry one can adjust the contours in such a way that $f(s)$, as defined in (B1) be treated in the same manner as $\xi(s)$ in (15) – see Figure 1 – viz.

$$N_f(T) = \frac{1}{2\pi i}\int_{c_1'+c_2} d\log f(s) = \frac{1}{\pi}\operatorname{Im}\int_{c_1+c_2'} d\log f(s). \qquad \text{(B2)}$$

Implications are also the same in that $\operatorname{Re}\int_{c_1'+c_2} d\log f(s) = 0$, and

$$\arg f(\tfrac{1}{2}+iT) = \arg\sqrt{f(\tfrac{1}{2}+\delta+iT)\cdot f(\tfrac{1}{2}-\delta+iT)}\,. \qquad \text{(B3)}$$

By inspection it's obvious that $\left|f(\tfrac{1}{2}+\delta+iT)\right| = \left|f^*(\tfrac{1}{2}+\delta+iT)\right| = \left|f(\tfrac{1}{2}+\delta-iT)\right|$, which shows the rather trivial result $\operatorname{Re}\int_{c_1'} d\log f(s) = 0$. Note also that if we take $\delta\to-\delta$

---

[20] It has been established that any off critical zero must be generated through $\varsigma(s)$ and not by $\Omega(s)$

when inputting (B1) into (B3) then $\left|f\left(\tfrac{1}{2}+\delta+iT\right)\right|=\left|f\left(\tfrac{1}{2}-\delta+iT\right)\right|$ and, through symmetry of the real component, one can easily verify that

$$\mathrm{Re}\int_{c_2} d\log f(s)=\log\left(\left|\frac{f\left(\tfrac{1}{2}+\delta+iT\right)}{f\left(\tfrac{1}{2}-\delta+iT\right)}\right|\right)=\log(1)=0 \qquad \text{(B4)}$$

For the imaginary component that counts zeros in (B2), one must be careful which zeros to consider in the context of the contour chosen. The fact that only *one* zero is to be considered for the half contour integral in (B2) while *two* are considered for the full contour integral (divided by 2) implies that only the zeros above the real axis are ever really considered – giving the exact same result if one was to ignore the complex conjugate zeros below the real axis and just consider the first two factors in (B1) – i.e. redefine the function, $f(s)=(s-\rho_\alpha)(s+\rho_\alpha^*+1)$ instead. One of course must account for all 4 zeros in matters concerning symmetry as the contour $c_1'=c_3+c_1$ requires a zero below the real axis to be present to support the zero picked up in c_3 via the transformation $s\to 1-s$ as considered in the main section of this paper; however, since the integral is unchanged from the original contour only 2 zeros need be included in context of any relative measures involving contour integration.

The idea is to distinguish cases where the zeros are inside or outside the fully (half) enclosed contour. This is measured in 2 dimensions. The first measures whether one is high enough in the plane to capture zeros: I. $T<t_\alpha$ and II. $T>t_\alpha$. The other dimension determines whether one is wide enough to capture the zeros: a) $\delta<\alpha$ and b) $\delta>\alpha$. In case I the height of the contour is less than the height of the zero, hence any zero will necessarily remain outside the contours regardless of whether a) $\delta<\alpha$ or b) $\delta>\alpha$. In case II, the zero is outside the contour for a) and inside for b); hence, by the Residue theorem – see Fig. 1

$$N_f(T)=\begin{cases}0, & \delta<\alpha,\ \text{case I a)}\\ 0, & \delta>\alpha,\ \text{case I b)}\\ 0, & \delta<\alpha,\ \text{case II a)}\\ 1, & \delta>\alpha,\ \text{case II b)}\end{cases}$$

The implication of this is that $\arg f\left(\tfrac{1}{2}+iT\right)=\pi N_f(T)$ is not single-valued and thus implicitly depends on the path of the chosen contour, as measured through the integral parameter $\delta:=\tfrac{1}{2}+\varepsilon$, just as was observed for $S(T)=\pi^{-1}\arg\varsigma\left(\tfrac{1}{2}+iT\right)$. In other words, for a given T dependency on the contour parameter ε (or δ) implies that, either (B1) and (B3) are no longer valid and we have a broken symmetry, or $f(s)$ *is* not well-defined at the points (poles) symmetric about the critical line – and across the cut line connecting them – where the zeros are the branch points, which no contour can pass. This is obviously the case for functions such as $\log f(s)$, $\sqrt{f(s)}$, $\arg f(s)$, and $\arg\sqrt{f(s)}$, etc. but is not expected to be an issue for the function itself as defined in (B1). However, in the context of (B1)-(B3) this symmetry must be adhered to as we are comparing a half contour with the full one, effectively taking the square root of the actual function and we then have a criterion for distinguishing between certain classes of function like that defined in (B1) with $\alpha > 0$ where a branch cut is required from one where a cut is not required to maintain symmetry – i.e. where $\alpha = 0$.

To illustrate take $\alpha > 0$ in (B1), then the function $\sqrt{f(s)}$ necessarily requires a cut line that connects the zeros, which are branch points, lying across the line $\mathrm{Re}(s) = \frac{1}{2}$ so as to ensure and maintain the single valued nature of the phase or argument.[21] In this way no contour may cross the branch cut, as its argument is discontinuous across that line. By construction of the contour integral in (B2) a cut line is required to make the function single valued when $\alpha > 0$, therefore one must know a priori the location of the zero ρ_α and maintain this parameter to ensure the single valued nature of the square root of the function. In the strictest sense the function defined in (B1) is not completely analytic by definition as is evident in the symmetry of (B2) and (B3) where the analytic nature of square root of the function is implied. This property is exactly what is observed when considering the existence of off critical zeros of the zeta function, generated *endogenously* by the fluctuating term in (19). From this point of view, consequences for RH being true or false are quite straightforward – e.g. in order that the function be single valued, the *multiplicity* (or residue) of the zeros (poles, in this context) must be at least *two*. This ensures that the function in question (B1) is single valued along with its square root in such a way that the symmetry in (B2) and (B3) is maintained. However, this doesn't address the multi-valued nature of the argument or *phase* of the function. Detailed discussions of consequences to RH being true or false in context of the holomorphic function (1) defined by Riemann are provided in the main section. The only way to make the argument of that function single valued, without crossing a branch cut, is to make $\alpha = 0$ – see below:

Consider what happens when $\alpha = 0$. First of all, at a given height T the integral (B2) does not cross any branch cut except across the critical line for the full contour integral. This is not an issue since the zeroes on the line do not lie on the contour; i.e. they are effectively fixed points and so remain smooth and analytic throughout the strip. Given that the zeros on the line are complex conjugates $\rho_\alpha = 1 - \rho_\alpha^*$, one is then able to unambiguously take the square root of (B1) and, therefore

$$h(s) := \sqrt{f(s)} = (s - \rho_\alpha)(s + \rho_\alpha - 1) = (s - \rho_\alpha)(s - \rho_\alpha^*), \qquad \text{(B5)}$$

which is essentially the same function as before affecting (B1) by a multiplicity of 2. Again, by symmetry of the integral in (B2) one only needs to consider the 1st factor in (B5). In this case no branch points or cut lines are crossed to make $\arg f(s)$ single valued below. It is therefore analytic and invariant with respect to δ – that is to say (B3) for both $f(s)$ and $\sqrt{f(s)}$, as defined in (B1) *is* valid. Therefore through symmetry considerations alone, the argument of the square root function $\arg h(s) = \arg\sqrt{f(s)} = \frac{1}{2}\arg f(s)$ is also single valued, symmetric, (piecewise) continuous and analytic, provided zeroes lie on the

[21] See e.g. Sect. 7.1 of Arfken, "Mathematical Methods for Physicists", 3rd Edition (Academic Press, 1985)

axis of symmetry – i.e. $\alpha = 0$. In fact this implies that any variable power law parameter will affect the argument (or *phase*) by a simple scale invariant factor, otherwise nothing has changed. This of course, also implies there are no branch cut singularities for $\log f(s)$ and $\log h(s)$ as well – a result which is probably surprising, if skipping Appendix I, as ordinarily a branch cut *is* required for the logarithm and square root if applied to simple complex functions.

For this simple case ($\alpha = 0$), it is an easy task to verify (B3) for $h(s) = \sqrt{f(s)}$. When looking at the left hand side one has – keeping in mind the half contour picks up only the zero above the real axis

$$\arg h(\tfrac{1}{2} + iT) = \operatorname{Im}\log\left[0_+ + i(T - t_\alpha)\right] = \begin{cases} 0, & T < t_\alpha \\ \pi/2, & T > t_\alpha \end{cases} \tag{B6}$$

For the right hand side of (B3) one has

$$\arg\sqrt{h(\tfrac{1}{2} + \delta + iT)\cdot h(\tfrac{1}{2} - \delta + iT)} = \frac{1}{2}\left\{\arctan\left(\frac{T - t_\alpha}{\delta}\right) + \arctan\left(\frac{T - t_\alpha}{-\delta}\right)\right\}. \tag{B7}$$

Using (A10, 11) of Appendix I and, carefully considering the cases I, II, a) & b) defined above – this result is independent of (or invariant to) δ and identical to (B6), which validates (B1) and (B3).

The class of functions considered here share characteristics with the simple function $f(s) := s(s-1)$ that cancels singularities of the functional representation in (4) and (6) in that it is also invariant under the transformation $s \to 1 - s$ and has zeros on either side of the critical line. However, these zeros explicitly cancel singularities inherent in the primordial functional representation given by Riemann – viz. $\xi(s)s^{-1}(s-1)^{-1}$ as evident in (4).

In precisely the same manner, we can show that it is invariant using the results from Appendix I:

$$\arg\left[(\tfrac{1}{2} + iT)(-\tfrac{1}{2} + iT)\right] = \arctan\left(\frac{T}{\tfrac{1}{2}}\right) + \arctan\left(\frac{T}{-\tfrac{1}{2}}\right) = \pi \tag{B8}$$

For the right hand side one has

$$\arg\sqrt{(\tfrac{1}{2} - \delta + iT)\cdot(-\tfrac{1}{2} - \delta + iT)\cdot(\tfrac{1}{2} + \delta + iT)\cdot(-\tfrac{1}{2} + \delta + iT)} = \tag{B9}$$

$$\frac{1}{2}\left\{\arctan\left(\frac{T}{\frac{1}{2}-\delta}\right)+\arctan\left(\frac{T}{-\frac{1}{2}-\delta}\right)+\arctan\left(\frac{T}{\frac{1}{2}+\delta}\right)+\arctan\left(\frac{T}{-\frac{1}{2}+\delta}\right)\right\}.$$

From (A10, 11) of Appendix I this gives a value of π for (B9), which verifies (B3) for the simple case $f(s) := s(s-1)$.

Although elementary, the class of functions considered in this appendix serve to illustrate that, while the function in (B2) is analytic in one context – i.e. through the Cauchy-Riemann conditions, symmetry considerations do not make it well-defined in the context of (B1) unless a multiplicity of 2 or branch cut singularity is accounted for. Since the symmetric function defined in (1), (4) and (6) is analytic in the entire complex plane – i.e. *holomorphic* – no branch cut singularity can be present. This implies that RH is necessarily true.

References

1. Riemann, B., Translated from *Ueber die Auzahl de Primzahlen unter einer gegebenn Grösse*, by H.M. Edwards – see *Gesammelte Werke*. Teubner, Leipzig 1892 (reprinted by Dover books, New York, 1953).
2. Hardy, G. H., *Sur les zeros de la function do Riemann*, C. R. Acad. Sci. Paris **158** (1914), 1012-1014.
3. Littlewood, J. E., *On the zeros of the Riemann Function*, Proc Cambridge Philos. Soc. **22** (1924), 295-318.
4. Hardy, G. H., & Littlewood, J. E., *The zeros of Riemann's zeta-function on the critical line*, Math. Z. **10** (1921), 283-317.
5. Bohr H. & Landau, E. *Ein Satz über Dirichletsche Riehen mit Anwendung*… Rend. Circ. Mat. Palermo **37** (1914) 269-272.
6. Selberg, A. *The zeta-function and the Riemann Hypothesis*, Proc. Skand. Math. Kongr 10th (1947) 187-200.
7. P. Erdös, 1949, *On a new method in elementary number theory which leads to an elementary proof of the prime number theorem*, Proc. Nat. Acad. Sci. U. S. A., **35**, 374-384; Selberg, A. *An elementary proof of the prime number theorem*, Ann. Of Math [2] **50** (1949), 305-313.
8. Turing, A. M., *Some calculations of the Riemann zeta function*, Proc. London Math. Soc. [3], **3** (1953), 99-117.
9. Siegel, C. L., über *Riemanns Nachlass* sur analytischen Zahlentheorie. Quellen Studien zur Geschichte der Math. Astron. und Phys. Abt. B: Studien **2** (1932).
10. Riemann, B. Unpublished papers, *Handscriftenabteilung Niedersächische Staatsund* Universitätsbibliothek, Göttigen.
11. von Mangoldt, H. Zu Riemann's Abhandlung, *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, J. Reine Angew. Math. **114** (1895), 255-305.
12. Hadamard, J. *Sur la distribution des zeros de la function* $\varsigma(s)$ *et ses consequences arithmétiques*. Bull. Soc. Math. France **24** (1896), 199-220.
13. de la Vallée Poussin, D.-J., *Recherches analytiques surla théorie des nombres* (premiere partie). Ann. Soc. Sci. Bruxelles [1], **20_2** (1896), 183-256.
14. Titchmarsh, E.C., *The theory of the Riemann zeta function*, 2nd ed. Rev. Heath-Brown, D. R. (Oxford University Press, 1986)

15. Hadamard, J., *Étude sur les Propiétés des function Entières et en Particulier d'une Fonction Considérée par Riemann*, J. Math. Pures Appl. [4] 9 (1893), 171-215.
16. Edwards, H. M. *Riemann's zeta function* (Dover, New York 1974)
17. Backlund, R. *Sur les zeros de la fonction* $\varsigma(s)$ *de Riemann*. C. R. Acad. Sci. Paris 158 (1914), 1979-1982.
18. Montgomery, H. L., *The pair correlation of zeros of the zeta function*, Analytic Number Theory, Proc. Symp. Pure Math Vol XXIV), 181-193, (Amer. Math. Soc., Providence RI, 1973).
19. Odlyzko, A. M., *The 10^22-nd zero of the Riemann zeta function*, Dynamical, Spectral, and Arithmetic Zeta Functions, M. van Frankenhuysen and M. L. Lapidus, eds., Amer. Math. Soc., Contemporary Math. series, no. 290, 2001, pp. 139-144. See also http://www.dtc.umn.edu/~odlyzko/zeta_tables/zeros3
20. Dyson, F. J., *A Brownian-motion model for the eigenvalues of a random matrix*, J. Math. Phys. 3 (1962), 1191-1198; (ibid.), *Statistical theory of the energy levels of complex systems*, J. Math. Phys. 3 (1962), 140-175.
21. Wigner, E., *On the statistical distribution of the widths and spacings of nuclear resonance levels*, Proc. Cambridge Phil. Soc. 47 (1951) 790-798.
22. P. Leboeuf, A. G. Monastra and O. Bohigas, *The Riemannium*, Regular and Chaotic Dynamics 6 (2001) 205-210.
23. Faddeev, L. D. and Pavlov, B. S., *Scattering theory and automorphic functions*, Seminar of Steklov Mathematical Institute of Leningrad 27 (1972) 161-193.
24. Lax, P. and Phillips, R. S., *Scattering Theory* (1st ed. 1967) revised edition – Pure and Applied Mathematics 26 (Academic Press, 1989).
25. S. Joffily, *Jost functions, prime numbers and Riemann zeta function*, math-phys/0303014.
26. Boos, H. E. and Korepin, V. E., *Quantum spin chains and Riemann zeta function with odd arguments*, J. Phys. A 34 (2001) 5311-5316; (ibid.) J. Phys. A 35 (2002) 4443-4452.
27. Braun, H.-B. and Loss, D., *Berry's Phase and quantum dynamics of ferromagnetic solitons*, Phys. Rev. B (cond. mat.) 53 (1996) 3237-3255 – see Appendix A.
28. Loss, D. and Maslov, D. L., *Spin stiffness of mesoscopic quantum anti-ferromagnets*, Phys. Rev. Lett. 74 (1995) 781-784.
29. Allen, S. and Loss, D. *Relationship between susceptibility and spin stiffness of mesoscopic quantum anti-ferromagnets*, Physica A 239 (1997) 47-52;
30. Allen, S. and Loss, D., to be published (2007).
31. Engel, H.-A. and Loss, D., *Conductance fluctuations in diffusive rings: Berry phase effects and criteria for adiabicity*, Phys. Rev. B 62 (2000) 10238.
32. Lee, T. D. and Yang, C. N., *Statistical theory of equations of state and phase transitions I. Theory of condensation*, Phys. Rev. 87 (1952) 404-409; (ibid.) II. *Lattice gas and Ising model*, 410-419.
33. Wolf, M., *Applications of statistical mechanics in number theory*, Physica A274 (1999) 149-157.
34. Julia, B. L., *Statistical theory of numbers*, from Number Theory and Physics, M. Waldschmidt, et al. (eds.), (Springer-Verlag, Berlin, Heidelberg 1990) pp. 276-293.
35. Wolf, M., *1/f noise in the distribution of prime numbers*, Physica 241 (1997) 493-499.
36. P. Bak, C. Tang, and K. Wiesenfeld, *Self-organised criticality*, Phys. Rev. A 38 (1988) 364-374.
37. Berry, M. V., & Keating, J. P., *The Riemann zeros and eigenvalue asymptotics*, SIAM Review 41 (1999), 236-266 – see referenced therein.
38. Bost, J.-B. and Connes, A., *Hecke Algebras, Type III factors and phase transitions with spontaneous symmetry breaking in number theory*, Selecta Math. (New series), 1 (1995) 411-457.
39. Julia, B. L., *Thermodynamic limit in number theory: Riemann-Beurling gases*, Physica A 203 (1994) 425-436.

40. Li, W. and Holste, D., *Universal 1/f noise, crossovers of scaling exponents and chromosome specific patterns of guanine-cytosine, content in DNA sequences of the human genome*, Phys. Rev. E (2005), 041910.
41. Hausdorff, J. M. and Peng, C.-K., *Multiscaled randomness: A possible source of 1/f noise in biology*, Phys. Rev. E 54 (1996), 2154.
42. Rosu, H. C., *Quantum Hamiltonians and prime numbers*, Mod. Phys. Lett. A 18 (2003).
43. Seahra, S. S. and Wesson, P. S., *Universes encircling 5-dimensional black holes*, J. Math. Phys. **44** (2003) 5664-5680
44. T. Fukui, S. S. Seahra, and P. S. Wesson, *Cosmological Implications of a non-separable 5D solution of the vacuum Einstein field equations*, J. Math. Phys. **42** (2001) 5195-5201.
45. Hu, B. L., *Semi-classical gravity and mesoscopic physics*, gr-qc/9511077
46. E. Calzetta, A. Campos, and E. Verdauguer, *Stochastic semi-classical cosmological models*, Phys. Rev. D **56** (1997) 2163-2172.
47. Phillips, N. G. and Hu, B. L., *Fluctuations of the vacuum energy density of quantum fields in curved space-time via generalized zeta functions*, Phys. Rev. D **55** (1997) 6123-6234.
48. Øksendal, B., *Stochastic differential equations*, 6th ed. Springer (2003).
49. Gutzwiller, M. C., *Classical quantization of a Hamiltonian with ergodic behaviour*, Phys. Rev. Lett. **45**(3) (1980) 150-153.
50. Berry, M. V., *Quantal phase factors accompanying adiabatic changes*, Proc. R. Soc. London A **392** (1984) 45-57.
51. Planat, M., *1/f frequency noise in a communication receiver and the Riemann Hypothesis*, in Noise, Oscillators and Algebraic Randomness: From Noise in Communication Systems to Number Theory (conference proceedings - Chapelle des Bois, France, April 5-10, 1999), M. Planat, ed. (Springer-Verlag, 2000).
52. Planat, M., *Huyghens, Bohr, Riemann and Galois: Phase-locking*, Int. J. Mod. Phys. B **20** (2006) 1833-1850.
53. Selberg, A., *Harmonic analysis and discontinuous groups in weakly symmetric Riemannian spaces with application to Dirichlet series*, Journal of the Indian Mathematical Society **20** (1956) 47-87.
54. Gutzwiller, M. C., *Chaos in Classical and Quantum Mechanics* (Springer-Verlag, 1990) p. 308.
55. Gutzwiller, M. C., *Quantum Chaos*, Scientific American, January 1992.
56. Hannay, J. H. and Ozorio De Almeida, A. M., *Period orbits and a correlation function for the density of states*, J. Phys. A **17** (1984) 3429-3440.
57. Pitkänen, M., *Riemann hypothesis and superconformal invariance*, math.GM/0102031.
58. Castro, C., *On p-adic stochastic dynamics, supersymmetry and the Riemann conjecture*, Chaos Solitons Fractals **14** (2002) 15-24.
59. C. Castro, A. Granik, and J. Mahecha, *On SUSY_QM, fractal strings and steps toward a proof of the Riemann hypothesis* – hep-th/0107266.
60. E. Elizalde, V. Moretti, and S. Zerbini, *On recent strategies proposed for proving the Riemann hypothesis*, Int. J. Mod. Phys. A **18** (2003) 2189.
61. Figueroa-O'Farrill, J. and Mohammedi, N., *Gauging the Wess-Zumino term of a sigma model with boundary*, JHEP **0508** (2005) 086.
62. Spector, D., *Supersymmetry and the Möbius inversion function*, Comm. Math. Phys. **127** (1990) 239-252.
63. Singh, S. & Pathria, R. K. *Analytical evaluation of a class of lattice sums in arbitrary dimensions*, J. Phys. A: Math. Gen. **22** (1989), 1883-1897
64. Allen, S. & Pathria, R.K., "On extended multi-dimensional Schlömilch series", J. Phys. A: Math. Gen. **28** (1995), 5757-5765. – see references therein, particularly the work of Ortner and Wagner.
65. Gradshteyn, I. M. and Ryzhik, I. M., *Table of Integrals, Series, and Products*, A. Jeffrey (ed.), Academic, New York, 1980.
66. Q. A Turchette, C. S. Wood, B. E. King, C. J. Myatt, D. Leibfried, W. N. Itano, C. Monroe, and D. J. Wineland, *Deterministic Entanglement of Two Trapped Ions*, Phys. Rev. Lett. **81** (1998), 3631-3634.

67. Bombin, H. and Martin-Delgado, M. A., *Entanglement Distillation Protocols and Number Theory* – quant-ph/0503013.
68. P. Binetruy, C. Deffayet, U. Ellewanger and D. Langlois, *Brane cosmological evolution in a bulk cosmological constant*, Phys. Lett B **277** (2000) 285-291.
69. Wesson, P. S., *Five-Dimensional Physics*, World Scientific, Singapore, 2006.
70. Connes, A., *Noncommutative geometry and the Riemann zeta function*, Mathematics: Frontiers and Perspectives 2000, V. Arnold et.al., eds. (2000) 35-55
71. Marcus du Sautoy, *The Music of the Primes: Searching to Solve the Greatest Mystery in Mathematics*, HarperCollins, 2003. ISBN 0-06-621070-4 – see also Matthew Watkins website http://secamlocal.ex.ac.uk/~mwatkins/zeta/dusautoy-keating.htm
72. John Derbyshire, *Prime Obsession: Bernhard Riemann and the Greatest Unsolved Problem in Mathematics*, Joseph Henry Press 2003. ISBN 0-30-908-549-7

Figure 1 – Integration contour(s) for Equation (7)

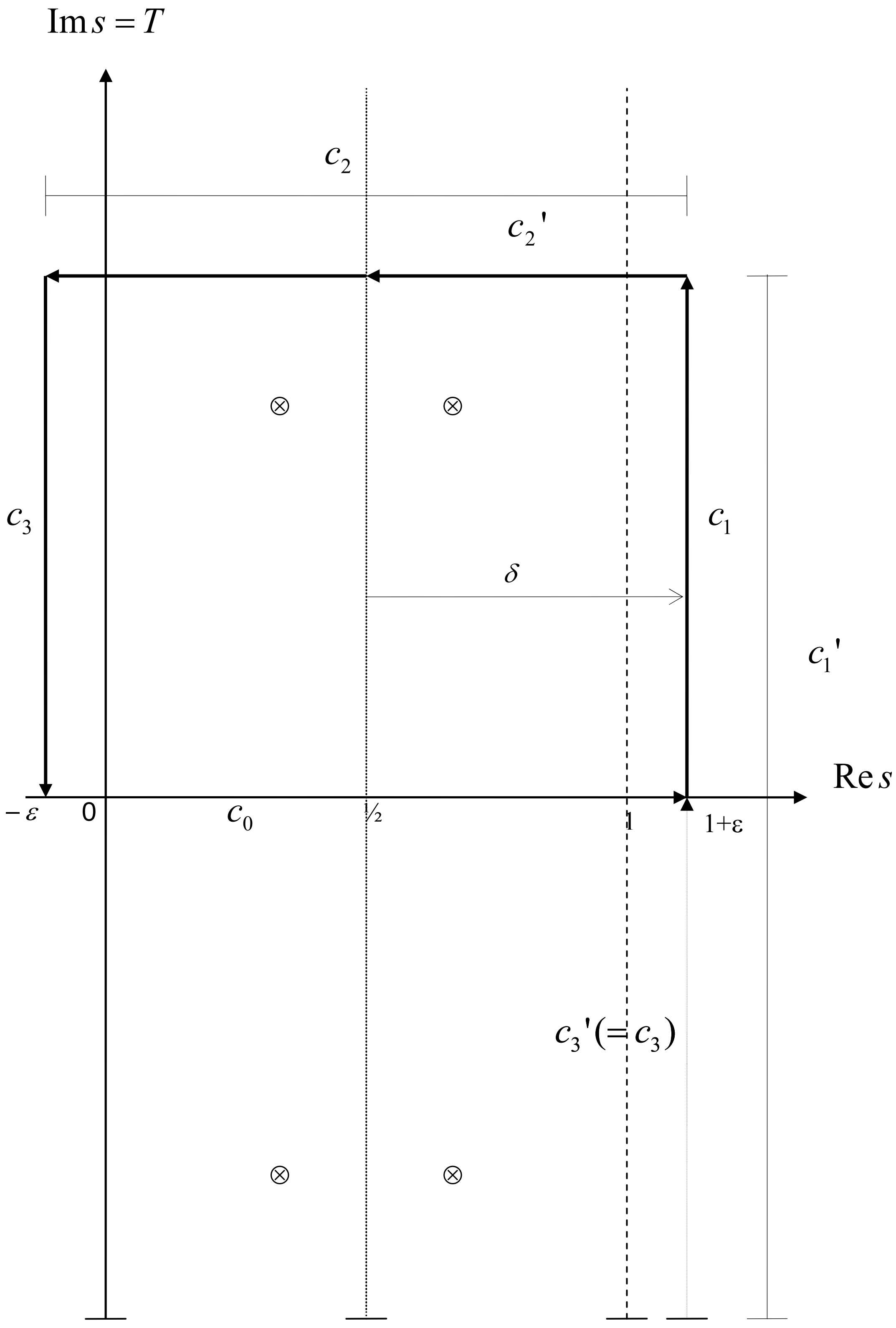

Figure 2