Acta Math. Univ. Comenianae
Vol. LXXII, 1(2003), pp. 1–13

A STRATEGY FOR PROVING RIEMANN HYPOTHESIS

M. PITKÄNEN

Abstract. A strategy for proving Riemann hypothesis is suggested. The vanishing of the Rieman Zeta reduces to an orthogonality condition for the eigenfunctions of a non-Hermitian operator D^+ having the zeros of Riemann Zeta as its eigenvalues. The construction of D^+ is inspired by the conviction that Riemann Zeta is associated with a physical system allowing conformal transformations as its symmetries. The eigenfunctions of D^+ are analogous to the so called coherent states and in general not orthogonal to each other. The states orthogonal to a vacuum state (which has a negative norm squared) correspond to the zeros of the Riemann Zeta. The induced metric in the space V of states which correspond to the zeros of the Riemann Zeta at the critical line $\mathrm{Re}[s] = 1/2$ is hermitian and hermiticity requirement actually implies Riemann hypothesis. Conformal invariance in the sense of gauge invariance allows only the states belonging to V. Riemann hypothesis follows also from a restricted form of a dynamical conformal invariance in V.

1. Introduction

The Riemann hypothesis [**6, 7**] states that the non-trivial zeros (as opposed to zeros at $s = -2n$, $n \geq 1$ integer) of Riemann Zeta function obtained by analytically continuing the function

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s} \tag{1}$$

from the region $\mathrm{Re}[s] > 1$ to the entire complex plane, lie on the line $\mathrm{Re}[s] = 1/2$. Hilbert and Polya conjectured a long time ago that the non-trivial zeroes of Riemann Zeta function could have spectral interpretation in terms of the eigenvalues of a suitable self-adjoint differential operator H such that the eigenvalues of this operator correspond to the imaginary parts of the nontrivial zeros $z = x + \mathrm{i}y$ of ζ. One can however consider a variant of this hypothesis stating that the eigenvalue spectrum of a non-hermitian operator D^+ contains the non-trivial zeros of ζ. The eigenstates in question are eigenstates of an annihilation operator type operator D^+ and analogous to the so called coherent states encountered in quantum physics [**4**]. In particular, the eigenfunctions are in general non-orthogonal and this is a quintessential element of the the proposed strategy of proof.

Received April 1, 2002.
2000 *Mathematics Subject Classification.* Primary 00-XX; Secondary 81-XX.
Key words and phrases. Riemann hypothesis, conformal invariance.

In the following an explicit operator having as its eigenvalues the non-trivial zeros of ζ is constructed.

a) The construction relies crucially on the interpretation of the vanishing of ζ as an orthogonality condition in a hermitian metric which is is a priori more general than Hilbert space inner product.

b) Second basic element is the scaling invariance motivated by the belief that ζ is associated with a physical system which has superconformal transformations [**3**] as its symmetries.

The core elements of the construction are following.

a) All complex numbers are candidates for the eigenvalues of D^+ (formal hermitian conjugate of D) and genuine eigenvalues are selected by the requirement that the condition $D^\dagger = D^+$ holds true in the set of the genuine eigenfunctions. This condition is equivalent with the hermiticity of the metric defined by a function proportional to ζ.

b) The eigenvalues turn out to consist of $z = 0$ and the non-trivial zeros of ζ and only the eigenfunctions corresponding to the zeros with $\mathrm{Re}[s] = 1/2$ define a subspace possessing a hermitian metric. The vanishing of ζ tells that the 'physical' positive norm eigenfunctions (in general *not* orthogonal to each other), are orthogonal to the 'unphysical' negative norm eigenfunction associated with the eigenvalue $z = 0$.

The proof of the Riemann hypothesis by reductio ad absurdum results if one assumes that the space V spanned by the states corresponding to the zeros of ζ inside the critical strip has a hermitian induced metric. Riemann hypothesis follows also from the requirement that the induced metric in the spaces subspaces V_s of V spanned by the states Ψ_s and $\Psi_{1-\overline{s}}$ does not possess negative eigenvalues. Conformal invariance in the sense of gauge invariance allows only the states belonging to V. Riemann hypothesis follows also from a restricted form of a dynamical conformal invariance in V.

2. Modified form of the Hilbert-Polya conjecture

One can modify the Hilbert-Polya conjecture by assuming scaling invariance and giving up the hermiticity of the Hilbert-Polya operator. This means introduction of the non-hermitian operators D^+ and D which are hermitian conjugates of each other such that D^+ has the nontrivial zeros of ζ as its complex eigenvalues

$$D^+\Psi = z\Psi. \tag{2}$$

The counterparts of the so called coherent states [**4**] are in question and the eigenfunctions of D^+ are not expected to be orthogonal in general. The following construction is based on the idea that D^+ also allows the eigenvalue $z = 0$ and that the vanishing of ζ at z expresses the orthogonality of the states with eigenvalue $z = x + iy \neq 0$ and the state with eigenvalue $z = 0$ which turns out to have a negative norm.

The trial

$$D = L_0 + V, \quad D^+ = -L_0 + V$$
$$L_0 = t\frac{d}{dt}, \qquad V = \frac{d\log(F)}{d(\log(t))} = t\frac{dF}{dt}\frac{1}{F} \tag{3}$$

is motivated by the requirement of invariance with respect to scalings $t \rightarrow \lambda t$ and $F \rightarrow \lambda F$. The range of variation for the variable t consists of non-negative real numbers $t \geq 0$. The scaling invariance implying conformal invariance (Virasoro generator L_0 represents scaling which plays a fundamental role in the superconformal theories [**3**]) is motivated by the belief that ζ codes for the physics of a quantum critical system having, not only supersymmetries [**1**], but also superconformal transformations as its basic symmetries (see the chapter "Riemann Hypothesis" of [**5**]).

3. Formal solution of the eigenvalue equation for operator D^+

One can formally solve the eigenvalue equation

$$D^+\Psi_z = \left[-t\frac{d}{dt} + t\,\frac{dF}{dt}\frac{1}{F}\right]\Psi_z = z\Psi_z. \tag{4}$$

for D^+ by factoring the eigenfunction to a product:

$$\Psi_z = f_z F. \tag{5}$$

The substitution into the eigenvalue equation gives

$$L_0 f_z = t\frac{d}{dt}f_z = -zf_z \tag{6}$$

allowing as its solution the functions

$$f_z(t) = t^z. \tag{7}$$

These functions are nothing but eigenfunctions of the scaling operator L_0 of the superconformal algebra analogous to the eigenstates of a translation operator. A priori all complex numbers z are candidates for the eigenvalues of D^+ and one must select the genuine eigenvalues by applying the requirement $D^\dagger = D^+$ in the space spanned by the genuine eigenfunctions.

It must be emphasized that Ψ_z is *not* an eigenfunction of D. Indeed, one has

$$D\Psi_z = -D^+\Psi_z + 2V\Psi_z = z\Psi_z + 2V\Psi_z. \tag{8}$$

This is in accordance with the analogy with the coherent states which are eigenstates of annihilation operator but not those of creation operator.

4. $D^+ = D^\dagger$ condition and hermitian form

The requirement that D^+ is indeed the hermitian conjugate of D implies that the hermitian form satisfies

$$\langle f|D^+g\rangle = \langle Df|g\rangle. \tag{9}$$

This condition implies

$$\langle \Psi_{z_1} | D^+ \Psi_{z_2} \rangle = \langle D\Psi_{z_1} | \Psi_{z_2} \rangle. \tag{10}$$

The first (not quite correct) guess is that the hermitian form is defined as an integral of the product $\overline{\Psi}_{z_1}\Psi_{z_2}$ of the eigenfunctions of the operator D over the non-negative real axis using a suitable integration measure. The hermitian form can be defined by continuing the integrand from the non-negative real axis to the entire complex t-plane and noticing that it has a cut along the non-negative real axis. This suggests the definition of the hermitian form, not as a mere integral over the non-negative real axis, but as a contour integral along curve C defined so that it encloses the non-negative real axis, that is C

a) traverses the non-negative real axis along the line $\mathrm{Im}[t] = 0_-$ from $t = \infty + \mathrm{i}0_-$ to $t = 0_+ + \mathrm{i}0_-$,

b) encircles the origin around a small circle from $t = 0_+ + \mathrm{i}0_-$ to $t = 0_+ + \mathrm{i}0_+$,

c) traverses the non-negative real axis along the line $\mathrm{Im}[t] = 0_+$ from $t = 0_+ + \mathrm{i}0_+$ to $t = \infty + \mathrm{i}0_+$.

Here $0_\pm$ signifies taking the limit $x = \pm\varepsilon$, $\varepsilon > 0$, $\varepsilon \to 0$.

C is the correct choice if the integrand defining the inner product approaches zero sufficiently fast at the limit $\mathrm{Re}[t] \to \infty$. Otherwise one must assume that the integration contour continues along the circle S_R of radius $R \to \infty$ back to $t = \infty + \mathrm{i}0_-$ to form a closed contour. It however turns out that this is not necessary. One can deform the integration contour rather freely: the only constraint is that the deformed integration contour does not cross over any cut or pole associated with the analytic continuation of the integrand from the non-negative real axis to the entire complex plane.

Scaling invariance dictates the form of the integration measure appearing in the hermitian form uniquely to be dt/t. The hermitian form thus obtained also makes possible to satisfy the crucial $D^+ = D^\dagger$ condition. The hermitian form is thus defined as

$$\langle \Psi_{z_1} | \Psi_{z_2} \rangle = -\frac{K}{2\pi i} \int_C \overline{\Psi_{z_1}} \Psi_{z_2} \frac{dt}{t}. \tag{11}$$

K is a real numerical constant which can be fixed by requiring that the states corresponding to zeros at the critical line have unit norm: with this choise the vacuum state corresponding to $z = 0$ has negative norm.

The possibility to deform the shape of C in wide limits realizes conformal invariance stating that the change of the shape of the integration contour induced by a conformal transformation, which is nonsingular inside the integration contour, leaves the value of the contour integral of an analytic function unchanged. This scaling invariant hermitian form is indeed a correct guess. By applying partial integration one can write

$$\langle \Psi_{z_1} | D^+ \Psi_{z_2} \rangle = \langle D\Psi_{z_1} | \Psi_{z_2} \rangle - \frac{K}{2\pi \mathrm{i}} \int_C dt \frac{d}{dt} \left[\overline{\Psi}_{z_1}(t) \Psi_{z_2}(t) \right]. \tag{12}$$

The integral of a total differential comes from the operator $L_0 = td/dt$ and must vanish. For a non-closed integration contour C the boundary terms from the partial integration could spoil the $D^+ = D^\dagger$ condition unless the eigenfunctions vanish at the end points of the integration contour ($t = \infty + i0_\pm$).

The explicit expression of the hermitian form is given by

$$\begin{aligned} \langle \Psi_{z_1} | \Psi_{z_2} \rangle &= -\frac{K}{2\pi i} \int_C \frac{dt}{t} F^2(t) t^{z_{12}} , \\ z_{12} &= \overline{z}_1 + z_2 . \end{aligned} \tag{13}$$

It must be emphasized that it is $\overline{\Psi}_{z_1}\Psi_{z_2}$ rather than eigenfunctions which is continued from the non-negative real axis to the complex t-plane: therefore one indeed obtains an analytic function as a result.

An essential role in the argument claimed to prove the Riemann hypothesis is played by the crossing symmetry

$$\langle \Psi_{z_1} | \Psi_{z_2} \rangle = \langle \Psi_0 | \Psi_{\overline{z}_1 + z_2} \rangle \tag{14}$$

of the hermitian form. This symmetry is analogous to the crossing symmetry of particle physics stating that the S-matrix is symmetric with respect to the replacement of the particles in the initial state with their antiparticles in the final state or vice versa [**4**].

The hermiticity of the hermitian form implies

$$\langle \Psi_{z_1} | \Psi_{z_2} \rangle = \overline{\langle \Psi_{z_2} | \Psi_{z_1} \rangle} . \tag{15}$$

This condition, which is *not* trivially satisfied, in fact determines the eigenvalue spectrum.

5. How to choose the function F?

The remaining task is to choose the function F in such a manner that the orthogonality conditions for the solutions Ψ_0 and Ψ_z reduce to the condition that ζ or some function proportional to ζ vanishes at the point $-z$. The definition of ζ based on analytical continuation performed by Riemann suggests how to proceed. Recall that the expression of ζ converging in the region $\mathrm{Re}[s] > 1$ following from the basic definition of ζ and elementary properties of Γ function [**7**] reads as

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{dt}{t} \frac{\exp(-t)}{[1 - \exp(-t)]} t^s . \tag{16}$$

One can analytically continue this expression to a function defined in the entire complex plane by noticing that the integrand is discontinuous along the cut extending from $t = 0$ to $t = \infty$. Following Riemann it is however more convenient to consider the discontinuity for a function obtained by multiplying the integrand with the factor

$$(-1)^s \equiv \exp(-i\pi s) .$$

The discontinuity $\mathrm{Disc}(f) \equiv f(t) - f(t\,\exp(i2\pi))$ of the resulting function is given by

$$\mathrm{Disc}\left[\frac{\exp(-t)}{[1 - \exp(-t)]} (-t)^{s-1} \right] = -2i \sin(\pi s) \frac{\exp(-t)}{[1 - \exp(-t)]} t^{s-1} . \tag{17}$$

The discontinuity vanishes at the limit $t \to 0$ for $\mathrm{Re}[s] > 1$. Hence one can define ζ by modifying the integration contour from the non-negative real axis to an integration contour C enclosing non-negative real axis defined in the previous section.

This amounts to writing the analytical continuation of $\zeta(s)$ in the form

$$-2\mathrm{i}\Gamma(s)\zeta(s)\sin(\pi s) = \int_C \frac{dt}{t}\frac{\exp(-t)}{[1-\exp(-t)]}(-t)^{s-1}. \tag{18}$$

This expression equals to $\zeta(s)$ for $\mathrm{Re}[s] > 1$ and defines $\zeta(s)$ in the entire complex plane since the integral around the origin eliminates the singularity.

The crucial observation is that the integrand on the righthand side of Eq. 18 has precisely the same general form as that appearing in the hermitian form defined in Eq. 13 defined using the same integration contour C. The integration measure is dt/t, the factor t^s is of the same form as the factor $t^{\overline{z}_1+z_2}$ appearing in the hermitian form, and the function $F^2(t)$ is given by

$$F^2(t) = \frac{\exp(-t)}{1-\exp(-t)}.$$

Therefore one can make the identification

$$F(t) = \left[\frac{\exp(-t)}{1-\exp(-t)}\right]^{1/2}. \tag{19}$$

Note that the argument of the square root is non-negative on the non-negative real axis and that $F(t)$ decays exponentially on the non-negative real axis and has $1/\sqrt{t}$ type singularity at origin. From this it follows that the eigenfunctions $\Psi_z(t)$ approach zero exponentially at the limit $\mathrm{Re}[t] \to \infty$ so that one can use the non-closed integration contour C.

With this assumption, the hermitian form reduces to the expression

$$\begin{aligned}\langle\Psi_{z_1}|\Psi_{z_2}\rangle &= -\frac{K}{2\pi\mathrm{i}}\int_C \frac{dt}{t}\frac{\exp(-t)}{[1-\exp(-t]}(-t)^{z_{12}} \\ &= \frac{K}{\pi}\sin(\pi z_{12})\Gamma(z_{12})\zeta(z_{12}).\end{aligned} \tag{20}$$

Recall that the definition $z_{12} = \overline{z}_1 + z_2$ is adopted. Thus the orthogonality of the eigenfunctions is equivalent to the vanishing of $\zeta(z_{12})$.

6. Study of the hermiticity condition

In order to derive information about the spectrum one must explicitely study what the statement that $D^\dagger$ is hermitian conjugate of D means. The defining equation is just the generalization of the equation

$$A^\dagger_{mn} = \overline{A}_{nm}. \tag{21}$$

defining the notion of hermiticity for matrices. Now indices m and n correspond to the eigenfunctions Ψ_{z_i}, and one obtains

$$\langle\Psi_{z_1}|D^+\Psi_{z_2}\rangle = z_2\langle\Psi_{z_1}|\Psi_{z_2}\rangle = \overline{\langle\Psi_{z_2}|D\Psi_{z_1}\rangle} = \overline{\langle D^+\Psi_{z_2}|\Psi_{z_1}\rangle} = z_2\overline{\langle\Psi_{z_2}|\Psi_{z_1}\rangle}.$$

Thus one has

$$\begin{aligned} G(z_{12}) &= \overline{G(z_{21})} = \overline{G(\overline{z}_{12})} \\ G(z_{12}) &\equiv \langle \Psi_{z_1} | \Psi_{z_2} \rangle . \end{aligned} \tag{22}$$

The condition states that the hermitian form defined by the contour integral is indeed hermitian. This is *not* trivially true. hermiticity condition obviously determines the spectrum of the eigenvalues of D^+.

To see the implications of the hermiticity condition, one must study the behaviour of the function $G(z_{12})$ under complex conjugation of both the argument and the value of the function itself. To achieve this one must write the integral

$$G(z_{12}) = -\frac{K}{2\pi \mathrm{i}} \int_C \frac{dt}{t} \frac{\exp(-t)}{[1-\exp(-t)]} (-t)^{z_{12}}$$

in a form from which one can easily deduce the behaviour of this function under complex conjugation. To achieve this, one must perform the change $t \to u =$ $= \log(\exp(-\mathrm{i}\pi)t)$ of the integration variable giving

$$G(z_{12}) = -\frac{K}{2\pi \mathrm{i}} \int_D du \frac{\exp(-\exp(u))}{[1-\exp(-(\exp(u)))]} \exp(z_{12}u). \tag{23}$$

Here D denotes the image of the integration contour C under $t \to u = \log(-t)$. D is a fork-like contour which
a) traverses the line $\mathrm{Im}[u] = \mathrm{i}\pi$ from $u = \infty + \mathrm{i}\pi$ to $u = -\infty + \mathrm{i}\pi$,
b) continues from $-\infty + \mathrm{i}\pi$ to $-\infty - \mathrm{i}\pi$ along the imaginary u-axis (it is easy to see that the contribution from this part of the contour vanishes),
c) traverses the real u-axis from $u = -\infty - \mathrm{i}\pi$ to $u = \infty - \mathrm{i}\pi$,

The integrand differs on the line $\mathrm{Im}[u] = \pm \mathrm{i}\pi$ from that on the line $\mathrm{Im}[u] = 0$ by the factor $\exp(\mp \mathrm{i}\pi z_{12})$ so that one can write $G(z_{12})$ as integral over real u-axis

$$G(z_{12}) = -\frac{K}{\pi} \sin(\pi z_{12}) \int_{-\infty}^{\infty} du \frac{\exp(-\exp(u))}{[1-\exp(-(\exp(u)))]} \exp(z_{12}u). \tag{24}$$

From this form the effect of the transformation $G(z) \to \overline{G(\overline{z})}$ can be deduced. Since the integral is along the real u-axis, complex conjugation amounts only to the replacement $z_{21} \to z_{12}$, and one has

$$\begin{aligned} \overline{G(\overline{z}_{12})} &= -\frac{\overline{K}}{\pi} \times \overline{\sin(\pi z_{21})} \int_{-\infty}^{\infty} du \frac{\exp(-\exp(u))}{[1-\exp(-(\exp(u)))]} \exp(z_{12}u) \\ &= \frac{\overline{K}}{K} \times \frac{\overline{\sin(\pi z_{21})}}{\sin(\pi z_{12})} G(z_{12}). \end{aligned} \tag{25}$$

Thus the hermiticity condition reduces to the condition

$$G(z_{12}) = \frac{\overline{K}}{K} \times \frac{\overline{\sin(\pi z_{21})}}{\sin(\pi z_{12})} \times G(z_{12}). \tag{26}$$

The reality of K guarantees that the diagonal matrix elements of the metric are real.

For non-diagonal matrix elements there are two manners to satisfy the hermiticity condition.

a) The condition

$$G(z_{12}) = 0 \tag{27}$$

is the only manner to satisfy the hermiticity condition for $x_1+x_2 \neq n$, $y_1-y_2 \neq 0$. This implies the vanishing of ζ:

$$\zeta(z_{12}) = 0 \text{ for } 0 < x_1 + x_2 < 1. \tag{28}$$

In particular, this condition must be true for $z_1 = 0$ and $z_2 = 1/2 + \mathrm{i}y$. Hence the physical states with the eigenvalue $z = 1/2+\mathrm{i}y$ must correspond to the zeros of ζ.

b) For the non-diagonal matrix elements of the metric the condition

$$\exp(\mathrm{i}\pi(x_1 + x_2)) = \pm 1 \tag{29}$$

guarantees the reality of $\sin(\pi z_{12})$ factors. This requires

$$x_1 + x_2 = n. \tag{30}$$

The highly non-trivial implication is that the the vacuum state Ψ_0 and the zeros of ζ at the critical line span a space having a hermitian but not necessarily positive definite metric. Note that for $x_1 = x_2 = n/2$, $n \neq 1$, the diagonal matrix elements of the metric vanish.

7. Various assumptions implying Riemann hypothesis

As found, the general strategy for proving the Riemann hypothesis, originally inspired by superconformal invariance, leads to the construction of a set of eigenstates for an operator D^+, which is effectively an annihilation operator acting in the space of complex-valued functions defined on the real half-line. Physically the states are analogous to coherent states and are not orthogonal to each other. The quantization of the eigenvalues for the operator D^+ follows from the requirement that the metric, which is defined by the integral defining the analytical continuation of ζ, and thus proportional to ζ ($\langle s_1, s_2\rangle \propto \zeta(\overline{s}_1 + s_2)$), is hermitian in the space of the physical states.

The nontrivial zeros of ζ are known to belong to the critical strip defined by $0 < \mathrm{Re}[s] < 1$. Indeed, the theorem of Hadamard and de la Vallee Poussin [**2**] states the non-vanishing of ζ on the line $\mathrm{Re}[s] = 1$. If s is a zero of ζ inside the critical strip, then also $1-\overline{s}$ as well as $\overline{s}$ and $1-s$ are zeros. Hilbert space inner product property is not required so that the eigenvalues of the metric tensor can be also negative. The problem is whether there could be also unphysical zeros of ζ outside the critical line $\mathrm{Re}[s] = 1/2$ but inside the critical strip $0 < \mathrm{Re}[s] < 1$.

Before continuing it is convenient to introduce some notations. Denote by V the subspace spanned by Ψ_s corresponding to the zeros s of ζ inside the critical strip, by V_{crit} the subspace corresponding to the zeros of ζ at the critical strip, and by V_s the space spanned by the states Ψ_s and $\Psi_{1-\overline{s}}$. The basic idea behind the following proposals is that the basic objects of study are the spaces V, V_{crit} and V_s.

7.1. How to restrict the metric to V?

One should somehow restrict the metric defined in the space spanned by the states Ψ_s labelled by a continuous complex eigenvalue s to the space V inside the critical strip spanned by a basis labelled by discrete eigenvalues. Very naively, one could try to do this by simply putting all other components of the metric to zero so that the states outside V correspond to gauge degrees of freedom. This is consistent with the interpretation of V as a coset space formed by identifying states which differ from each other by the addition of a superposition of states which do not correspond to zeros of ζ.

An more elegant manner to realize the restriction of the metric to V is to Fourier expand states in the basis labelled by a complex number s and define the metric in V using double Fourier integral over the complex plane and Dirac delta function restricting the labels of both states to the set of zeros inside the critical strip:

$$\begin{aligned}
\langle \Psi^{1)} | \Psi^{2)} \rangle &= \int d\mu(s_1) \int d\mu(s_2) \overline{\Psi}^{1)}_{s_1} \Psi^{2)}_{s_2} G(s_2 + \overline{s}_1) \delta(\zeta(s_1)) \delta(\zeta(s_2)) \\
&= \sum_{\zeta(s_1)=0, \zeta(s_2)=0} \overline{\Psi}^{1)}_{s_1} \Psi^{2)}_{s_2} G(s_2 + \overline{s}_1) \frac{1}{\sqrt{\det(s_2)\det(\overline{s}_1)}} , \\
d\mu(s) &= ds d\overline{s}, \qquad \det(s) = \frac{\partial(\mathrm{Re}\,[\zeta(s)]\,, \mathrm{Im}\,[\zeta(s)])}{\partial(\mathrm{Re}\,[s]\,, \mathrm{Im}\,[s])} .
\end{aligned} \tag{31}$$

Here the integrations are over the critical strip. $\det(s)$ is the Jacobian for the map $s \to \zeta(s)$ at s. The appearence of the determinants might be crucial for the absence of negative norm states. The result means that the metric G_V in V effectively reduces to a product

$$\begin{aligned}
G_V &= \overline{D} G D , \\
D(s_i, s_j) &= D(s_i) \delta(s_i, s_j) , \\
\overline{D}(s_i, s_j) &= D(\overline{s_i}) \delta(s_i, s_j) \\
D(s) &= \frac{1}{\sqrt{\det(s)}} .
\end{aligned} \tag{32}$$

In the sequel the metric G will be called reduced metric whereas G_V will be called the full metric. In fact, the symmetry $D(s) = D(\overline{s})$ holds true by the basic symmetries of ζ so that one has $D = \overline{D}$ and $G_V = DGD$. This means that Schwartz inequalities for the eigen states of D^+ are not affected in the replacement of G_V with G. The two metrics can be in fact transformed to each other by a mere scaling of the eigenstates and are in this sense equivalent.

7.2. Riemann hypothesis from the hermicity of the metric in V

The mere requirement that the metric is hermitian in V implies the Riemann hypothesis. This can be seen in the simplest manner as follows. Besides the zeros at the critical line $\mathrm{Re}[s] = 1/2$ also the symmetrically related zeros inside critical strip have positive norm squared but they do not have hermitian inner products with the states at the critical line unless one assumes that the inner product vanishes. The assumption that the inner products between the states at critical line and outside it vanish, implies additional zeros of ζ and, by repeating the argument again and again, one can fill the entire critical interval $(0, 1)$ with the zeros of ζ so that a reductio ad absurdum proof for the Riemann hypothesis results. Thus the metric gives for the states corresponding to the zeros of the Riemann Zeta at the critical line a special status as what might be called physical states.

It should be noticed that the states in V_s and $V_{\overline{s}}$ have non-hermitian inner products for $Re[s] \neq 1/2$ unless these inner products vanish: for $Re[s] > 1/2$ this however implies that ζ has a zero for $Re[s] > 1$.

7.3. Riemann hypothesis from the requirement that the metric in V_s does not possess negative eigenvalues

The requirement that the induced metric in the space V_s does not possess negative eigenvalues implies also Riemann hypothesis. The explicit expression for the norm of a $Re[s] = 1/2$ state with respect to the full metric G_V^{ind} reads as

$$\begin{aligned} G_V^{\mathrm{ind}}(1/2 + iy_n, 1/2 + iy_n) &= D^2(1/2 + iy)G^{\mathrm{ind}}(1/2 + iy_n, 1/2 + iy_n), \\ G^{\mathrm{ind}}(1/2 + iy_n, 1/2 + iy_n) &= -\frac{K}{\pi}\sin(\pi)\Gamma(1)\zeta(1). \end{aligned} \tag{33}$$

Here G^{ind} is the metric in V_s induced from the reduced metric G. This expression involves formally a product of vanishing and infinite factors and the value of expression must be defined as a limit by taking in $Im[z_{12}]$ to zero. The requirement that the norm squared defined by G^{ind} equals to one fixes the value of K:

$$K = -\frac{\pi}{\sin(\pi)\zeta(1)} = 1. \tag{34}$$

The components G^{ind} in V_s are given by

$$\begin{aligned} G^{\mathrm{ind}}(s, s) &= -\frac{\sin(2\pi Re[s])\Gamma(2Re[s])\zeta(2Re[s])}{\pi}, \\ G^{\mathrm{ind}}(1 - \overline{s}, 1 - \overline{s}) &= -\frac{\sin(2\pi(1 - Re[s]))\Gamma(2 - 2Re[s])\zeta(2(1 - [Re[s]))}{\pi}, \\ G^{\mathrm{ind}}(s, 1 - \overline{s}) &= G^{\mathrm{ind}}(1 - \overline{s}, s) = 1. \end{aligned} \tag{35}$$

The determinant of the metric G_V^{ind} induced from the full metric reduces to the product

$$\det(G_V^{\mathrm{ind}}) = D^2(s))D^2(1 - \overline{s}) \times Det(G^{\mathrm{ind}}). \tag{36}$$

Since the first factor is positive definite, it suffices to study the determinant of G^{ind}. At the limit $Re[s] = 1/2$ G^{ind} formally reduces to

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}.$$

This reflects the fact that the states Ψ_s and $\Psi_{1-\bar{s}}$ are identical. The actual metric is of course positive definite. For $\mathrm{Re}[s]=0$ the G^{ind} is of the form

$$\begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} .$$

The determinant of G^{ind} is negative so that the eigenvalues of both the full metric and reduced metric are of opposite sign. The eigenvalues for G^{ind} are given by $(-1\pm\sqrt{5})/2$.

The determinant of G^{ind} in V_s as a function of $\mathrm{Re}[s]$ is symmetric with respect to $\mathrm{Re}[s]=1/2$, equals to -1 at the end points $\mathrm{Re}[s]=0$ and $\mathrm{Re}[s]=1$, and vanishes at $\mathrm{Re}[s]=1/2$. Numerical calculation shows that the sign of the determinant of G^{ind} inside the interval $(0,1)$ is negative for $\mathrm{Re}[s]\neq 1/2$. Thus the diagonalized form of the induced metric has the signature $(1,-1)$ except at the limit $\mathrm{Re}[s]=1/2$, when the signature formally reduces to $(1,0)$. Thus Riemann hypothesis follows if one can show that the metric induced to V_s does not allow physical states with a negative norm squared. This requirement is physically very natural. It must be however emphasized that it is not clear whether the restriction of the metric to V_{crit} has only non-negative eigenvalues.

7.4. Riemann hypothesis and conformal invariance

The basic strategy for proving Riemann hypothesis has been based on the attempt to reduce Riemann hypothesis to invariance under conformal algebra or some subalgebra of the conformal algebra in V or V_s. That this kind of algebra should act as a gauge symmetry associated with ζ is very natural idea since conformal invariance is in a well-defined sense the basic symmetry group of complex analysis.

Consider now one particular strategy based on conformal invariance in the space of the eigenstates of D^+.

a) The conformal generators are realized as operators

$$L_z \quad = \quad t^z D^+ \tag{37}$$

act in the eigenspace of D^+ and obey the standard conformal algebra without central extension [**3**]. D^+ itself corresponds to the conformal generator L_0 acting as a scaling. Conformal generators obviously act as dynamical symmetries transforming eigenstates of D^+ to each other. What is new is that now conformal weights z have all possible complex values unlike in the standard case in which only integer values are possible. The vacuum state Ψ_0 having negative norm squared is annihilated by the conformal algebra so that the states orthogonal to it (non-trivial zeros of ζ inside the critical strip) form naturally another subspace which should be conformally invariant in some sense. Conformal algebra could act as gauge algebra and some subalgebra of the conformal algebra could act as a dynamical symmetry.

b) The definition of the metric in V involves in an essential manner the mapping $s\rightarrow\zeta(s)$. This suggests that one should define the gauge action of the conformal algebra as

$$\begin{aligned} \Psi_s &\rightarrow \Psi_{\zeta(s)} \rightarrow L_z\Psi_{\zeta(s)} = \zeta_s\Psi_{\zeta(s)+z} \\ &\rightarrow \zeta_s\Psi_{\zeta^{-1}(\zeta(s)+z)} . \end{aligned} \tag{38}$$

Clearly, the action involves a map of the conformal weight s to $\zeta(s)$, the action of the conformal algebra to $\zeta(s)$, and the mapping of the transformed conformal weight $z+\zeta(s)$ back to the complex plane by the inverse of ζ. The inverse image is in general non-unique but in case of V this does not matter since the action annihilates automatically all states in V. Thus conformal algebra indeed acts as a gauge symmetry. This symmetry does not however force Riemann hypothesis.

c) One can also study the action of the conformal algebra or its suitable subalgebra in V_s as a dynamical (as opposed to gauge) symmetry realized as

$$\Psi_s \rightarrow L_z\Psi_s = s\Psi_{s+z} . \tag{39}$$

The states Ψ_s and $\Psi_{1-\overline{s}}$ in V_s have nonvanishing norms and are obtained from each other by the conformal generators $L_{1-2\mathrm{Re}[s]}$ and $L_{2\mathrm{Re}[s]-1}$. For $\mathrm{Re}[s]\neq 1/2$ the generators $L_{1-2\mathrm{Re}[s]}$, $L_{2\mathrm{Re}[s]-1}$, and L_0 generate $SL(2,R)$ algebra which is non-compact and generates infinite number of states from the states of V_s. At the critical line this algebra reduces to the abelian algebra spanned by L_0. The requirement that the algebra naturally associated with V_s is a dynamical symmetry and thus generates only zeros of ζ leads to the conlusion that all points $s+n(1-2\mathrm{Re}[s])$, n integer, must be zeros of ζ. Clearly, $\mathrm{Re}[s]=1/2$ is the only possibility so that Riemann hypothesis follows. In this case the dynamical symmetry indeed reduces to a gauge symmetry.

There is clearly a connection with the argument based on the requirement that the induced metric in V_s does not possess negative eigenvalues. Since $SL(2,R)$ algebra acts as the isometries of the induced metric for the zeros having $\mathrm{Re}[s]\neq 1/2$, the signature of the induced metric must be $(1,-1)$.

d) One could even try to prove that the entire subalgebra of the conformal algebra spanned by the generators with conformal weights $n(1-2\mathrm{Re}[s])$ acts as a symmetry generating new zeros of ζ. If this holds, $\mathrm{Re}[s]=1/2$ is the only possibility so that Riemann hypothesis follows. In this case the dynamical conformal symmetry indeed reduces to a gauge symmetry.

Since $L_{1-2\mathrm{Re}[s]}$ acts as an infinitesimal isometry leaving the matrix element $\langle\Psi_0|\Psi_s\rangle=0$ invariant, one can in spirit of Lie group theory argue that also the exponentiated transformations $\exp(tL_{1-2\mathrm{Re}[s]})$ have the same property for all values of t. The exponential action leaves Ψ_0 invariant and generates from Ψ_s a superposition of states with conformal weights $s+n(1-2\mathrm{Re}[s])$, which all must be orthogonal to Ψ_0 since t is arbitrary. Therefore $\mathrm{Re}[s]=1/2$ is the only possibility.

7.5. Conclusions

To sum up, it seems that a promising approach for proving Riemann hypothesis is to demonstrate that the metric induced to V is hermitian. The hermiticity property reduces to the requirement that the dynamical conformal algebra naturally spanned by the states in V_s reduces to the abelian algebra defined by

$L_0 = D^+$. If infinitesimal isometries for the matrix elements $\langle\Psi_0|\Psi_s\rangle = 0$ generated by $L_{1-2\mathrm{Re}[s]}$ can be exponentiated, Riemann hypothesis follows. Contrary to the original physics motivated expectations, the metric defined by G (nor G_V) does not seem to be positive definite in V unless one poses additional conditions to the allowed superpositions of the eigenstates. This is due to the exponential increase of the moduli of the matrix elements $G(1/2+iy_1, 1/2+iy_2)$ for large values of $|y_1 - y_2|$ implying the failure of the Schwartz inequality when $y_{12} = y_1 - y_2$ is large.

Acknowledgment. I want to express my deep gratitude to Dr. Matthew Watkins for providing me with information about Riemann Zeta and for generous help, in particular for reading the earlier versions of the work and pointing out several inaccuracies and errors. I am also grateful for Prof. Masud Chaichian and Doc. Claus Montonen for encouraging comments and help.

References

1. Berry M. V. and Keating J. P., *Supersymmetry and Trace Formulae: Chaos and Disorder*, (I. V. Lerner, J. P. Keating, and D. E. Khmelnitskii eds.) Kluwer, New York, 1999, 355–367.
2. Edwards H. M., *Riemann's Zeta Function*, Academic Press, New York, 1974, 68–76.
3. Itzykson C., Saleur H. and Zuber J-B. (eds.), *Conformal Invariance and Applications to Statistical Mechanics* Word Scientific, Singapore, 1988.
4. Itzykson C. and Zuber J-B., *Quantum Field Theory*, McGraw-Hill Inc., New York, 1980, p. 549.
5. Pitkänen M., *Topological Geometrodynamics* (online book at `http://www.physics.helsinki.fi/~matpitka/tgd.html`, 2001.
6. Riemann B., *Ueber die Anzahl der Primzahlen unter einer gegebenen Grösse*, Monat. der Königl. Preuss. Akad. der Wissen. zu Berlin aus der Jahre 1859, 671–680; also, Gesammelte math. Werke und wissensch. Nachlass, 2. Aufl. 1892, 145–155.
7. Titchmarch E. C., *The Theory of the Riemann Zeta Function*, 2nd ed. revised by R. D. Heath-Brown, Oxford Univ. Press, Oxford, 1986.

M. Pitkänen, Department of Physical Sciences, High Energy Physics Division, PL 64, FIN-00014, University of Helsinki, Finland., *current address*: Kadermonkatu 16, 10900, Hanko, Finland., *e-mail*: `matpitka@rock.helsinki.fi`,
`http:www.physics.helsinki.fi/~matpitka`