Irregular Scattering, Number Theory, and Statistical Mechanics

Andreas Knauf*

Contribution to the Max Born Symposium, Wrocław, 15–17 Sept., 1993

1 Introduction

In this paper we present results on the relations between wave scattering in the so-called modular domain, analytic number theory, and classical statistical mechanics. The motion in the modular domain is an example of so-called arithmetic chaos. In this paper we shortly review results from [5, 6, 7], but concentrate on previously unpublished results.

The *modular group* $\Gamma \equiv \mathrm{PSL}(2,\mathbb{Z}) := \mathrm{SL}(2,\mathbb{Z})/\{\pm 1\!\!1\}$ acts on the *upper half plane* $\mathbb{H} := \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$ by *Möbius transformations*:

$$\hat{M}(z) := \frac{az+b}{cz+d},$$

$M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$ being a determinant one matrix with integer entries.

The *modular domain* (see Fig. 1)

$$\mathbb{F} := \{z \in \mathbb{H} \mid -\tfrac{1}{2} < \mathrm{Re}(z) < \tfrac{1}{2}, |z| > 1\}$$

is a *fundamental domain* of that group action, that is, every $z \in \mathbb{H}$ can be carried by $\hat{M} \in \Gamma$ to a point $\hat{M}(z) \in \bar{\mathbb{F}}$, and no point of $\mathbb{F}$ can be carried to another point of $\mathbb{F}$ in such a way.

The upper half plane is endowed with the Riemannian metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2}, \quad z \equiv x + iy \in \mathbb{H} \tag{1}$$

*Technische Universität, Fachbereich 3 - Mathematik, MA 7–2, Strasse des 17. Juni 135, D-10623 Berlin, Germany. e-mail: knauf@math.tu-berlin.de

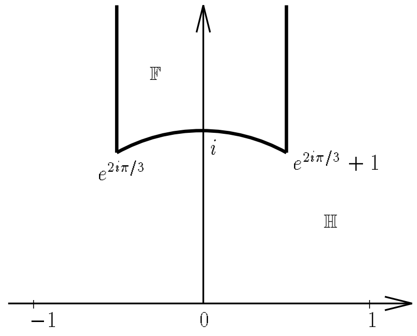

Figure 1: The modular domain $\mathbb{F}$

of curvature -1, which is invariant under arbitrary Möbius transformations.

As can be seen from (1), the Laplace-Beltrami operator Δ on $\mathbb{H}$ has the form $\Delta = y^2(\partial_x^2 + \partial_y^2)$. We shall analyze the non-Euclidean wave equation

$$\frac{\partial^2 u}{\partial t^2} = \left(\Delta + \frac{1}{4}\right) u \tag{2}$$

restricted to the modular domain $\mathbb{F}$. In other words, we consider initial data for (2) which are *automorphic functions* $f : \mathbb{H} \to \mathbb{C}$, $f \circ \hat{M} = f$ for $\hat{M} \in \Gamma$. Then by Γ-invariance of Δ the time-t solutions are automorphic, too.

In [3] Fadeev and Pavlov worked out the scattering theory on $\mathbb{F}$, that is, the automorphic solutions of (2) which come from and return to the cusp at $z = i\infty$. They showed that the scattering matrix $\mathcal{S}$ has the form

$$\mathcal{S}(\omega) = \frac{\Gamma(\frac{1}{2})\Gamma(i\omega)}{\Gamma(\frac{1}{2}+i\omega)} \cdot \frac{\zeta(2i\omega)}{\zeta(1+2i\omega)}, \tag{3}$$

ζ being the Riemann zeta function.

In [8, 9] Lax and Phillips analyzed scattering in the modular domain and related questions in the context of their abstract scattering theory.

The underlying classical system being geodesic motion, it should not come as a surprise that (3) has a semiclassical interpretation in terms of geodesics coming from and going to the cusp. This interpretation is presented in Sect. 2. What may be more astonishing is that semiclassics can be described in terms of a classical spin chain. In Sect. 3 we describe the interaction between the spins in that chain and present some results on its thermodynamic behaviour.

In the last section, we discuss the question whether that thermodynamic description can lead to new results in scattering theory and in number theory.

2 Scattering in the Modular Domain

In this section we give a short review of the derivation of the scattering matrix (3) and interpret it semiclassically.

The x-independent solutions of the wave equation (2) are of the form

$$u(z,t) = \sqrt{y}\cdot\left(u_-(\ln(y)+t) + u_+(\ln(y)-t)\right).$$

In particular, one has plane waves of frequency ω

$$u_\omega(z,t) = \sqrt{y}\cdot\left(c_- e^{i\omega(\ln(y)+t)} + c_+ e^{i\omega(-\ln(y)+t)}\right). \tag{4}$$

As remarked in Balazs and Voros [2], $\ln(y) = \int_{z=i}^{z=iy} ds$ is the classical action and $\sqrt{y}$ the square root of the invariant density, so that (4) is a semiclassical (WKB) wave.

We now construct automorphic generalized solutions of the wave equation by starting from $h_\omega(z) := y^{\frac12 + i\omega}$. h_ω is already invariant under the subgroup $\Gamma_\infty \subset \Gamma$ of integer translations $z \mapsto z+n$. So we only sum over the right cosets $\Gamma_\infty \setminus \Gamma$ and obtain the *Eisenstein series*

$$e(z,\omega) := \sum_{\hat{M}\in\Gamma_\infty\setminus\Gamma} h_\omega(\hat{M}(z)) \tag{5}$$

which is known to converge for frequencies ω with $\mathrm{Im}(\omega) < -\frac12$. By construction (5) is automorphic w.r.t. $z \in \mathbb{H}$. The scattering matrix $\mathcal{S}(\omega)$ is defined with the help of the zeroth x-Fourier mode

$$e^{(0)}(y,\omega) := \int_{-\frac12}^{\frac12} e(x+iy,\omega)dx, \qquad y>0,$$

namely

$$e^{(0)}(y,\omega) = y^{\frac12 + i\omega} + \mathcal{S}(\omega) y^{\frac12 - i\omega}. \tag{6}$$

If $M = \binom{a\,b}{c\,d} \in \mathrm{SL}(2,\mathbb{Z})$ is not in the coset space of the identity, then $c \neq 0$ and we can assume $c>0$ because we are dealing with $\Gamma = \mathrm{SL}(2,\mathbb{Z})/\{\pm 1\!\!1\}$. Then the other elements of the right coset of M are of the form $\binom{a-nc\;\;b-nd}{c\qquad d}$, $n\in\mathbb{Z}$, so that there is a unique representative with $0<a\le c$. The greatest common divisor $(c,d)=1$, since $ad-bc=1$. If, on the other hand, we are given $c>0$ and d with $(c,d)=1$, then there is a unique $M=\binom{a\,b}{c\,d}\in\mathrm{SL}(2,\mathbb{Z})$ with $0<a\le c$. Furthermore,

$$h_\omega(\hat{M}(z)) = \left(\mathrm{Im}\left(\frac{az+b}{cz+d}\right)\right)^{\frac12+i\omega} = \frac{y^{\frac12+i\omega}}{[(cx+d)^2+c^2y^2]^{\frac12+i\omega}},$$

so that one need not determine the integers a and b. Thus one gets

$$e(z,\omega) = y^{\frac12+i\omega}\left(1+\sum_{c\in\mathbb{N}}\sum_{d\in\mathbb{Z},(c,d)=1}[(cx+d)^2+c^2y^2]^{-\frac12-i\omega}\right)$$

and

$$\begin{aligned} e^{(0)}(y,\omega) &= y^{\frac12+i\omega}\left(1+\sum_{c\in\mathbb{N}}\sum_{d\in\mathbb{Z},(c,d)=1}\int_{-\frac12}^{\frac12}[(cx+d)^2+c^2y^2]^{-\frac12-i\omega}dx\right)\\ &= y^{\frac12+i\omega}\left(1+\sum_{c\in\mathbb{N}}\sum_{1\le d\le c,\,(c,d)=1}\int_{-\infty}^{\infty}[(cx+d)^2+c^2y^2]^{-\frac12-i\omega}dx\right)\\ &= y^{\frac12+i\omega} + y^{\frac12-i\omega}\frac{\Gamma(\frac12)\Gamma(i\omega)}{\Gamma(\frac12+i\omega)}\sum_{c\in\mathbb{N}}\frac{\varphi(c)}{c^{1+2i\omega}} \end{aligned} \tag{7}$$

with the Euler totient $\varphi(c) = \#\{d\in\{1,\ldots,c\} \mid (c,d)=1\}$.

By (6) this implies the Fadeev-Pavlov formula (3) for the scattering matrix $\mathcal{S}(\omega)$, since the quotient $\zeta(s-1)/\zeta(s)$ has the Dirichlet series

$$\frac{\zeta(s-1)}{\zeta(s)} = \sum_{n\in\mathbb{N}}\frac{\varphi(n)}{n^s}.$$

Now we will give a semiclassical geometrical interpretation of the Fadeev-Pavlov formula in terms of geodesics.

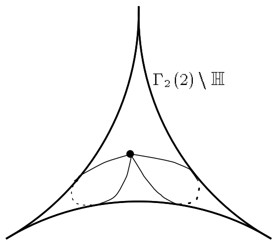

Figure 2: An embedding of the surface $\Gamma_2(2) \setminus \mathbb{H}$ in $\mathbb{R}^3$, with generators of its fundamental group

When we identify corresponding points on the boundary of $\bar{\mathbb{F}}$, we obtain the Riemannian surface $\Gamma \setminus \mathbb{H}$ which is of finite volume. But the action of the modular group Γ on $\mathbb{H}$ is not free: the point $i \in \mathbb{H}$ is a fixed point of the involutive transformation $z \mapsto -1/z$, and similarly $\hat{S}(e^{2\pi i/3}+1) = e^{2\pi i/3}+1$ for $S = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z})$, $\hat{S}$ generating an order three subgroup. If we try to endow $\Gamma \setminus \mathbb{H}$ with the metric coming from $\mathbb{H}$, this leads to conical singularities.

If we want to get rid of these problems, we may consider a smooth six-fold covering surface of $\Gamma \setminus \mathbb{H}$ which has the form $\Gamma_2(2) \setminus \mathbb{H}$, where $\Gamma_2(n) \subset \Gamma$ denotes the subgroup whose elements have matrix representations M of the form $M = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod n$. $\Gamma_2(2)$ is of index 6 in Γ and is freely generated by two generators with the matrices $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$.

$\Gamma_2(2)$ acts freely on $\mathbb{H}$ so that $\Gamma_2(2) \setminus \mathbb{H}$ is a smooth Riemannian manifold of area 2π and curvature -1. It has three cusps (say, $i\infty$, $\hat{S}(i\infty)$ and $\hat{S}^{-1}(i\infty)$), and it is homeomorphic to a sphere with three points deleted.

In Figure 2 we show $\Gamma_2(2) \setminus \mathbb{H}$ together with representatives of the two generators of its fundamental group. Although the embedding of the surface in $\mathbb{R}^3$ shown in Fig. 2 is not isometric (probably such an embedding does not exist), it is relatively easy to visualize geodesic motion on it.

There is another interesting aspect concerning $\Gamma_2(2) \setminus \mathbb{H}$. Namely, its fundamental group, a free group on two generators, is the same as that of a certain Riemann surface that was used in [4] in order to describe the motion

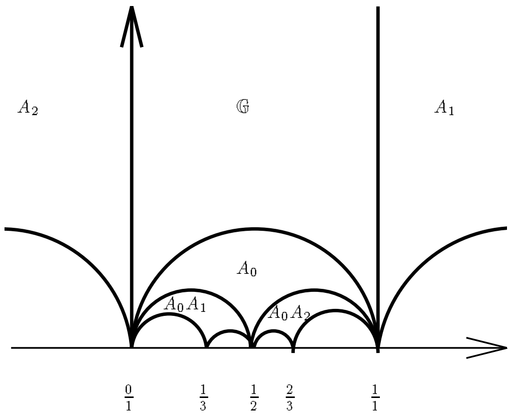

Figure 3: The fundamental domain $\mathbb{G}$ of the Farey tesselation

of a particle in the field of three nuclei. It is thus an open question whether parts of the analysis given in this paper carry over to the more physical situation of planar scattering by Coulombic potentials.

However, it turns out that for our purposes it is more useful to consider still another tesselation of the upper half plane, starting with the fundamental domain

$$\mathbb{G} := \{z \in \mathbb{H} \mid 0 < \mathrm{Re}(z) < 1, |z - \tfrac{1}{2}| > \tfrac{1}{2}\},$$

see Fig. 3.

$\mathbb{G}$ has three times the volume of $\mathbb{F}$ and the Möbius transformation $\hat{S}$ leads to a cyclic permutation of the three cusps at $i\infty$, 0 and 1.

$\mathbb{G}$ is the fundamental domain of the subgroup of Γ which is generated by the order two transformations $\hat{A}_0, \hat{A}_1, \hat{A}_2 \in \Gamma$ with matrices

$$A_0 := \begin{pmatrix} 1 & -1 \\ 2 & -1 \end{pmatrix}, \quad A_1 := SA_0S^{-1} = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix} \text{ and } A_2 := S^2A_0S^{-2} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Because of its connection with Farey fractions C. Series christened the corresponding tesselation the Farey tesselation [10].

Proposition 1 *Every $M \in \mathrm{SL}(2,\mathbb{Z})$ can be uniquely written in the form*

$$M = \pm A_{\pi(1)} \cdot \ldots \cdot A_{\pi(k)} \cdot S^j \tag{8}$$

with $\pi(n) \in \{0,1,2\}$, $\pi(n) \neq \pi(n+1)$, and $j \in \{0,1,2\}$ (For $k=0$ this notation means $M = \pm S^j$).

In the Appendix we give an algorithm for obtaining the above decomposition and, using that algorithm, uniqueness follows.

Corollary 2 *For $M \in \mathrm{SL}(2,\mathbb{Z})$ the domain $\hat{M}(\mathbb{G})$ is contained in*

0. $\{z \in \mathbb{H} \mid 0 < \mathrm{Re}(z) < 1\}$ *iff* $M = \pm S^j$ *or* $\pi(1) = 0$

1. $\{z \in \mathbb{H} \mid 1 < \mathrm{Re}(z)\}$ *iff* $\pi(1) = 1$

2. $\{z \in \mathbb{H} \mid \mathrm{Re}(z) < 0\}$ *iff* $\pi(1) = 2$

in the decomposition (8).

Proof. Since $\mathbb{G}$ is bounded by three geodesics ending in the cusps at ∞, 0 and 1, it suffices to observe that $\hat{A}_m(I_l) \subset I_m$, $m \neq l$, for the intervals $I_0 :=]0,1[$, $I_1 :=]1,\infty[$ and $I_2 :=]-\infty,0[$ between these points.

Then the proof follows by induction in $k \in \mathbb{N}$ from Proposition 1. □

From Prop. 1 and Cor. 2 it follows that there is a one-to-one correspondence between translates $\hat{M}(\mathbb{G}) \subset \{z \in \mathbb{H} \mid 0 < \mathrm{Re}(z) < 1\}$ and matrices in

$$\mathcal{U} := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \mid 0 \leq a \leq c, 0 \leq b \leq d \right\} \cup \{\mathbb{1}\}.$$

On the other hand, denoting by

$$\mathcal{U}_{\mathrm{LP}} := \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) \mid 1 \leq a \leq c, 1 \leq d \leq c \right\}$$

the class of matrices over which the summation in (7) is performed, we have the isomorphism

$$\mathcal{U}_{\mathrm{LP}} \to \mathcal{U}, \qquad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a-b & b \\ c-d & d \end{pmatrix}.$$

Therefore, we may rewrite the sum appearing in (7)

$$\begin{aligned} \sum_{c \in \mathbb{N}} \sum_{1 \leq d \leq c,\, (c,d)=1} c^{-1-2i\omega} &= \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}} (c+d)^{-1-2i\omega} \\ &= \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}} \exp(-2\ln(c+d) \cdot (\tfrac{1}{2} + i\omega)). \end{aligned} \tag{9}$$

This substitution has a simple semiclassical interpretation.

To every $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}$ we associate the vertical geodesic with real part $(a+b)/(c+d)$ coming from the middle of the three cusps of $\hat{M}(\mathbb{G})$, see Fig. 4 a).

The Ford circle (see, e.g., [1]) in $\mathbb{H}$ with center at $\frac{a+b}{c+d} + \frac{i}{2(c+d)^2}$ and radius $\frac{1}{2}(c+d)^{-2}$ is the image of the horizontal line $\mathrm{Im}(z) = 1$ under the Möbius transformation $z \mapsto \hat{M}\hat{S}^2(z)$, $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$. Equivalently, it is the image of the Ford circle $|z - (1+i/2)| = \frac{1}{2}$ under the transformation $z \mapsto \hat{M}(z)$. Thus it is a curve of constant phase w.r.t. the partial wave originating in $\frac{a+b}{c+d}$. Its vertical metric distance to that line equals $2\ln(c+d)$, see Fig. 4 b).

This means that (9) is a sum over partial waves associated to the geodesics going to the cusp. The factors $\exp(-2\ln(c+d) \cdot (\frac{1}{2} + i\omega))$ are coming from the time delays of these partial waves relative to the partial wave emanating at $z = 1$.

Moreover, Prop. 1 suggests a partition of these partial waves into families of 2^k members. This leads us to classical statistical mechanics.

3 The Number-Theoretical Spin Chain

In [5] we analysed the spin chain with k spins and configuration space $\mathbf{G}_k := \mathbb{Z}_2^k$, $\mathbb{Z}_2 = \{0,1\}$, whose canonical energy function $\mathbf{H}_k^C : \mathbf{G}_k \to \mathbb{R}$ is uniquely defined as follows: $\mathbf{H}_k^C := \ln(\mathbf{h}_k^C)$ with $\mathbf{h}_0^C := 1$, $\mathbf{h}_{k+1}^C(\sigma_1, \ldots, \sigma_k, 0) := \mathbf{h}_k^C(\sigma_1, \ldots, \sigma_k)$ and $\mathbf{h}_{k+1}^C(\sigma_1, \ldots, \sigma_k, 1) := \mathbf{h}_k^C(\sigma_1, \ldots, \sigma_k) + \mathbf{h}_k^C(1-\sigma_1, \ldots, 1-\sigma_k)$.

Considering the set $\mathcal{U}_k \subset \mathcal{U}$ containing the 2^k matrices of the form $M = \pm A_{\pi(1)} \cdot \ldots \cdot A_{\pi(n)} \cdot S^j$ with $n \leq k$ and $\pi(1) = 0$ if $n > 0$, we have

$$\sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}_k} (c+d)^{-s} = \sum_{\sigma \in \mathbf{G}_k} \left(\mathbf{h}_k^C(\sigma)\right)^{-s} \equiv \sum_{\sigma \in \mathbf{G}_k} \exp\left(-s \cdot \mathbf{H}_k^C(\sigma)\right).$$

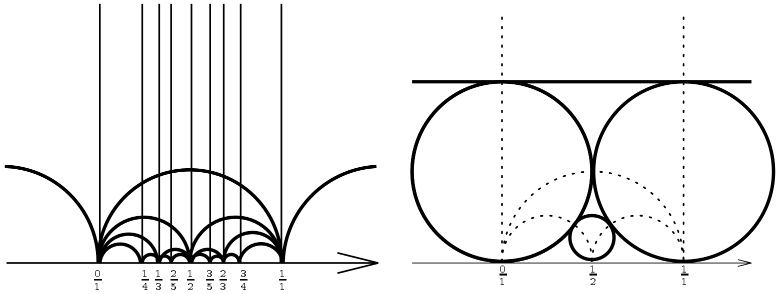

Figure 4: a) Geodesics related to the Farey tesselation. b) The Ford circles

Thus (9) can be formally interpreted as the partition function of an infinite spin chain for inverse temperature $s = \frac{1}{2} + i\omega$.

In [5] and [7] it was shown that this interpretation is not only formal. The canonical interaction coefficients $j_k^C(t)$, $t = (t_1, \dots, t_k) \in \mathbb{G}_k^* \cong \mathbb{G}_k$ are given by the negative Fourier transform of the canonical energy function $\mathbf{H}_k^C$:

$$j_k^C(t) := -2^{-k} \sum_{\sigma \in \mathbf{G}_k} \mathbf{H}_k^C(\sigma) \cdot (-1)^{\sigma \cdot t}.$$

Then $\mathbf{H}_k^C(\sigma) = -\sum_{t \in \mathbf{G}_k} j_k^C(t) \cdot (-1)^{\sigma \cdot t}$.

In [5] it was shown that the interaction is asymptotically translation invariant. The interaction is ferromagnetic, that is, $j_k^C(t) \geq 0$ for $t \neq 0$. The individual coefficients decay exponentially w.r.t. the distance of the spins involved in the interaction, but multi-body interactions play an important role.

The effective interaction decays like the squared inverse distance of the spins. It is well-known that for spin chains a faster polynomial decay of the interaction leads to a thermodynamic behaviour without phase transitions. In our case, however, a phase transition occurs, as shown in [6]. In [7] it was shown that for $\beta > 0$ the thermodynamic limit $F(\beta) := \lim_{k\to\infty} F_k(\beta)$ of the free energy per particle $F_k(\beta) := -\frac{1}{\beta \cdot k} \ln\left(Z_k(\beta)\right)$ with partition function $Z_k(s) := \sum_{\sigma \in \mathbf{G}_k} \exp(-s \cdot \mathbf{H}_k^C(\sigma))$ exists. Since for $\mathrm{Re}(s) > 2$ the partition

Figure 5: a) Two point correlation function. b) Inverse correlation length

function of the infinite chain is finite, i.e.

$$\lim_{k\to\infty} Z_k(s) = Z(s) := \frac{\zeta(s-1)}{\zeta(s)},$$

it is clear that $F(\beta) = 0$ in the low temperature range $\beta > 2$. Moreover, the mean magnetization per spin

$$M_k(\beta) := \frac{1}{k \cdot Z_k(\beta)} \sum_{\sigma \in \mathbf{G}_k} \left(\sum_{i=1}^{k} (1 - 2\sigma_i) \right) \exp(-\beta \cdot \mathbf{H}_k^C(\sigma))$$

has a thermodynamic limit $M(\beta)$ with $M(\beta) = 1$ for $\beta > 2$. Thus for low temperatures, due to the long range of the interaction, the system is in a frozen state.

On the other hand, numerical and analytical arguments suggest a phase transition of Thouless type at $\beta_{\mathrm{cr}} = 2$, the mean magnetization jumping to $M(\beta) = 0$ for $\beta < 2$ and the energy per spin increasing sharply.

Due to the ferromagnetic character of the interaction all correlation functions are positive.

In Fig. 5 a) we show a Monte Carlo calculation of the two point correlation function for a chain of $k = 900$ spins, averaged over all spins of a fixed distance. For $\beta < 2$ the correlations seem to decay exponentially w.r.t. the distance.

In Fig. 5b) the inverse correlation length is plotted as a function of β, together with a least square linear fit. The data are compatible with a phase transition at $\beta_{\mathrm{cr}} = 2$ and a critical exponent $\nu = 1$.

4 The Unreasonable Effectiveness of Statistical Mechanics

In spite of much constant effort, the behaviour of the Riemann zeta function $\zeta(s)$ in the critical strip $0 < \mathrm{Re}(s) < 1$ is not well understood. Because of the functional equation $\zeta(1-s) = 2(2\pi)^{-s}\Gamma(s)\cos(\pi s/2)\zeta(s)$ the values for arguments s with $\mathrm{Re}(s) < \frac{1}{2}$ are determined by the values for $\mathrm{Re}(s) > \frac{1}{2}$ and, according to the Riemann Hypothesis, all ζ function zeroes with non-vanishing imaginary part should lie on the critical line $\mathrm{Re}(s) = \frac{1}{2}$. Both the Dirichlet series and the Euler product representations are not convergent for $\mathrm{Re}(s) \leq 1$ so that one does not have a very direct method for calculating the values of the zeta function in the critical strip.

Thus the question arises whether the statistical mechanics approach leads to a method for calculating these values. $Z(s) = \zeta(s-1)/\zeta(s)$ has a pole at $s = 2$ which on the level of statistical mechanics probably gives rise to the above-mentioned phase transition. Thus it seems hopeless to find a convergent approximation to $Z(s)$ valid for $\mathrm{Re}(s) > 3/2$. But this is not necessarily so.

In the theory of Dirichlet series the notion of Bohr equivalence is important. Two Dirichlet series $\sum_{n=1}^{\infty} a(n) \cdot n^{-s}$ and $\sum_{n=1}^{\infty} b(n) \cdot n^{-s}$ are Bohr equivalent iff there exists a completely multiplicative function $f : \mathbb{N} \to \mathbb{C}$ with $b(n) = f(n) \cdot a(n)$, $n \in \mathbb{N}$, and $|f(p)| = 1$ whenever $a(n) \neq 0$ and p is a prime divisor of n.

In our case $\tilde{Z}(s) := \sum_{n=1}^{\infty} \lambda(n) \cdot \varphi(n) \cdot n^{-s}$ with the Liouville function $\lambda(\prod_{p_i \text{ prime}} p_i^{\alpha_i}) := (-1)^{\sum_i \alpha_i}$ is equivalent to the partition function $Z(s) = \sum_{n=1}^{\infty} \varphi(n) \cdot n^{-s}$. For $\mathrm{Re}(s) > 2$ one has

$$\tilde{Z}(s) = \prod_{p \text{ prime}} \frac{1+p^{-s}}{1+p^{1-s}} = \frac{\zeta(s) \cdot \zeta(2(s-1))}{\zeta(s-1) \cdot \zeta(2s)} = \frac{Z(2s) \cdot Z(2s-1)}{Z(s)},$$

so that the analytical continuation of $\tilde{Z}(s)$ has a zero instead of a pole at $s = 2$.

Writing $Z_k(s) = \sum_{n=1}^{\infty} \varphi_k(n) \cdot n^{-s}$ with $\varphi_k(n) := \#\{\sigma \in \mathbf{G}_k \mid \mathbf{h}_k^C(\sigma) = n\}$, $\tilde{Z}(s)$ is uniformly approximated in any half-plane $\mathrm{Re}(s) \geq 2 + \varepsilon$ by

$$\tilde{Z}_k(s) := \sum_{n=1}^{\infty} \lambda(n) \cdot \varphi_k(n) \cdot n^{-s}.$$

The question is whether one can find larger half-planes of convergence.

Because of the Riemann zeroes, $\tilde{Z}_k(s)$ has poles on the shifted critical line $\mathrm{Re}(s) = 3/2$ so that one cannot expect convergence for s with $\mathrm{Re}(s) \leq 3/2$.

In absence of a mathematical theory, we compared the numerical values of $\tilde{Z}_k(s)$ with $\tilde{Z}(s)$. Fig. 6 a) and b) show the contour plots of $|\tilde{Z}(s)|$ and $|\tilde{Z}_{25}(s)|$ in the window $1.5 \leq \mathrm{Re}(s) \leq 2.5$, $6 \leq \mathrm{Im}(s) \leq 12$.

Fig. 6 c) and d) show contour plots of the absolute values of the functions $\hat{Z}_m(s) := \sum_{n=1}^{m} \lambda(n) \cdot \varphi(n) \cdot n^{-s}$ for $m = 10\,000$ and $m = 250\,000$. In the first case $\hat{Z}_m$ has approximately the same number $\sum_{n=1}^{m} \varphi(n) \sim \frac{3}{\pi^2} m^2$ of terms n^{-s} as $\tilde{Z}_{25}$ (2^{25} terms). Yet, due to oscillations near the shifted critical line, the numerical accuracy of the approximation is much worse, and this accuracy does not increase much even for $m = 250\,000$. A similar behaviour is found for other values of $\mathrm{Im}(s)$.

A An Algorithm for decomposing $\mathrm{SL}(2,\mathbb{Z})$ Elements

In this Appendix we describe an algorithm for decomposing $M \in \mathrm{SL}(2,\mathbb{Z})$ in the form $M = \pm A_{\pi(1)} \cdot \ldots \cdot A_{\pi(k)} \cdot S^j$ described in Prop. 1, using the map

$$D : \mathrm{SL}(2,\mathbb{Z}) \to \mathbb{N}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \max(|a|,|b|,|c|,|d|,|a+b|,|c+d|).$$

Obviously we have $D(MS) = D(M)$.

Lemma 3 *The subset of $M \in \mathrm{SL}(2,\mathbb{Z})$ with $D(M) = 1$ equals $\{\pm 1, \pm S, \pm S^2, \pm A_2, \pm A_2 S, \pm A_2 S^2\}$.*

Proof. By inspection. □

If $D(M) = 1$, then the maximum in the definition of $D(M)$ is attained at least twice. To the contrary, for $D(M) > 1$ one has

Lemma 4 *If $D(M) > 1$, then only one element of the list $(|a|,|b|,|c|,|d|,|a+b|,|c+d|)$ equals $D(M)$.*

Proof. We prove the statement by considering all pairs in the list:

Figure 6: Contour plots of the absolute values of a) $\tilde{Z}(s)$ b) $\tilde{Z}_{25}(s)$ c) $\hat{Z}_{10\,000}(s)$ d) $\hat{Z}_{250\,000}(s)$

1. The greatest common divisor of the following pairs in the list equals one: $(a,b) = (b,c) = (c,d) = (d,a) = (a,a+b) = (b,a+b) = (c,c+d) = (d,c+d) = (a+b,c+d) = 1$. If the two partners are equal, their absolute value must be one.

2. If $|a| = |d| > 1$ then $|bc| = |ad-1| \geq |ad| - 1 = (a+1)(a-1) > 0$ which implies $\max(|b|,|c|) > |a|$ so that $|a| < D(M)$. The case $|b| = |c|$ is similar.

3. The pairs $(|a|,|c+d|)$, $(|b|,|c+d|)$, $(|c|,|a+b|)$ and $(|d|,|a+b|)$ can be transformed to the pairs of case 2) by right multiplication with $S^{\pm 1}$, remembering that $D(MS^{\pm 1}) = D(M)$. $\Box$

In order to decompose M we multiply M from the *right* by the unique A_i with $D(MA_i) < D(M)$ until $D(M) = 1$. By Lemma 3 this leads to a decomposition of the form $M = \pm S^l \cdot A_{\rho(1)} \cdot \ldots \cdot A_{\rho(k)}$. Then we transform this decomposition to the form of Prop. 1, using the relation $A_i = S^i A_0 S^{-i}$.

If $D(M) > 1$ by Lemma 4 we have the three disjoint cases

0. $D(M) = \max(|a|,|c|)$

1. $D(M) = \max(|b|,|d|)$

2. $D(M) = \max(|a+b|,|c+d|)$

which are cyclically transformed into each other by the transformation $M \mapsto MS$. So we need only treat one of them, say 2). In that case we assume $D(M) = |a+b|$, the case $D(M) = |c+d|$ being similar. Then $ab > 0$ and $cd \geq 0$ (since $abcd = bc(bc+1) \geq 0$) so that $D(MA_2) = \max(|a|,|b|,|c|,|d|,|a-b|,|c-d|) < D(M)$, whereas $D(MA_0) = \max(|a+2b|,|c+2d|) > D(M)$ and $D(MA_1) = \max(|2a+b|,|2c+d|) > D(M)$.

So in case 2) a multiplication of M with A_2 leads to a decrease of $D(M)$. By symmetry, we multiply in the case i) with A_i. This shows uniqueness of our algorithm and also uniqueness of the decomposition (8).

References

[1] Apostol, T.M.: Modular Functions and Dirichlet Series in Number Theory. Graduate Texts in Mathematics 41, New York: Springer 1990

[2] Balazs, N.L., Voros, A.: Chaos on the Pseudosphere. Physics Reports **143**, 109–240 (1986)

[3] Fadeev, L.D., Pavlov, B.S.: Scattering Theory and Automorphic Functions. Proc. Steklov Inst. Math. **27**, 161–193 (1972). English translation: J. Sov. Math. **3**, 522–548 (1975)

[4] Klein, M., Knauf, A.: Classical Planar Scattering by Coulombic Potentials. Lecture Notes in Physics m 13. Berlin, Heidelberg, New York: Springer 1993

[5] Knauf, A.: On a Ferromagnetic Spin Chain. Commun. Math. Phys. **153**, 77–115 (1993)

[6] Knauf, A.: Phases of the Number-Theoretical Spin Chain. *To appear in* Journal of Statistical Physics

[7] Knauf, A.: On a Ferromagnetic Spin Chain. Part II: Thermodynamic Limit. *To appear in* Journal of Mathematical Physics

[8] Lax, P.D., Phillips, R.S.: Scattering Theory for Automorphic Functions. Annals of Mathematics Studies 87, New Jersey: Princeton University Press 1976

[9] Lax, P.D., Phillips, R.S.: Scattering Theory for Automorphic Functions. Bulletin of the AMS (New Series) **2**, 261–295 (1980)

[10] Series, C.: The Modular Surface and Continued Fractions. J. London Math. Soc. **31**, 69–80 (1985)