Number theory, dynamical systems and statistical mechanics

Andreas Knauf*

May 1998

Abstract

In these lecture notes connections between the Riemann zeta function, motion in the modular domain and systems of statistical mechanics are presented.

1 Introduction

Counting was the earliest mathematical activity. Number theory thus was among the first subjects of mathematics. It was shown by Euclid that every integer $n \in \mathbb{N}$ has a unique factorization

$$n = \prod_{p \in \mathbb{P}} p^{\alpha_p} \tag{1}$$

in terms of the primes $\mathbb{P} \subset \mathbb{N}$, and a class of rings like $\mathbb{Z}$ sharing this property thus bears his name.

To some extent the beauty of number theory seems to be related to the contradiction between the simplicity of the integers and the complicated structure of the primes, their building blocks. This has always attracted people. A non-representative example is the following, from a book by Sacks: He describes a dialogue between twins:

> "John would say a number — a six-figure number. Michael would catch the number, nod, smile and seem to savour it. Then he, in turn, would say another six-figure number, and now it was John who received, and appreciated it richly. They looked, at first, like two connoisseurs wine-tasting, sharing rare tastes, rare appreciations."

*Max-Planck-Institute for Mathematics in the Sciences, Inselstr. 22–26, D-04103 Leipzig, Germany. e-mail: knauf@mis.mpg.de

After having recognized with the help of a table that these numbers are prime, Sacks returns the favour, when they meet next time, by presenting an eight-digit prime number. Some minutes later John answers with a nine-digit integer...

The twins described in the book *The man who mistook his wive for a hat* by the neuro-psychologist Oliver Sacks were inhabitants of a mental hospital. They were unable to perform even simple arithmetic operations. According to their own description, they *saw* the landscape of integers, which is indeed very bizarre.

Maybe I should quote a source which is nearer to the way mathematics is done at a Max-Planck-Institute. In his inaugural lecture at Bonn University, Don Zagier argues:

> "There are two facts about the distribution of prime numbers of which I hope to convince you so overwhelmingly that they will be permanently engraved in your hearts.
>
> The first is that, despite their simple definition and role as the building blocks of the natural numbers, the prime numbers belong to the most arbitrary and ornery objects studied by mathematicians: they grow like weeds among the natural numbers, seeming to obey no other law than that of chance, and nobody can predict where the next one will sprout.
>
> The second fact is even more astonishing, for it states just the opposite: that the prime numbers exhibit stunning regularity, that there are laws governing their behaviour, and that they obey these laws with almost military precision."[56]

This series of lecture notes is aimed at presenting a current attempt to use ideas from mathematical physics, more specifically from the ergodic theory of dynamical systems and equilibrium statistical mechanics, to solve problems in number theory.

The group of people having contributed to this attempt includes M. Berry, J.-B. Bost, A. Connes, P. Contucci, P. Cvitanović, M. Feigenbaum, F. Guerra, D. Harari, B. Julia, J. Keating, P. Kleban, O.E. Lanford, E. Leichtnam, A. Lopes, D. Mayer, A. Özlük, M. Pollicott, I. Procaccia, L. Ruedin, D. Spector, T. Tél and myself.

One basic observation in this attempt, and the one on which I will concentrate, is that *scattering in the modular domain* of $\mathrm{PSL}(2,\mathbb{Z})$ may be analyzed by using the *thermodynamic formalism*. The quantum scattering amplitude (which is basically a quotient of Riemann zeta functions) can be written in terms of the relative time delay of the geodesics coming from and going to the cusp.

Finally, that time delay on the set of scattering geodesics may be considered as the energy function of a spin chain, and the scattering amplitude turns out to be the partition function of that system.

The spin chain shows a phase transition, and its interaction function (the negative Fourier transform of the energy function) turns out to be positive.

This is of some interest because of the analogy with partition functions of ferromagnetic Ising systems, whose zeroes all lie on a line (Lee-Yang Theorem).

At this moment the above attempt has not been successful in deriving new number-theoretical results, so that its importance for number theory naturally remains controversial.

On the other hand, there are now numerous applications of results from number theory to physical problems. To mention just a few,

- p-adic analysis plays a role in string theory [7]
- a key element in understanding the growth of plants like sunflowers is the modular domain [32]
- some properties of information networks are optimal if they have the form of so-called Ramanujan graphs [50]
- finally, the Riemann zeta function ist the mother of all zeta functions, including the ones considered in the theory of dynamical systems [4, 44, 49].

As I shall deal with p-adic analysis, the modular domain and Ramanujan graphs, these lecture notes might be of some use even if I cannot convince you that there is a relevant leftward arrow in

$$\text{Number Theory} \iff \text{Physics} .$$

I begin by reviewing some elementary properties of the Riemann zeta function. Instead of giving detailed proofs, I shall indicate the main ideas and refer for details to the introductory literature.

Acknowledgement. I thank G. Rudolph and E. Zeidler for the opportunity to give this series of lectures in the 'Oberseminar Mathematische Physik'.

2 The Riemann Zeta Function

A central fact supporting Zagier's second observation is the

Prime Number Theorem. The number $\pi(x) := |\mathbb{P} \cap [0, x]|$ of primes not exceeding x is asymptotic to

$$\pi(x) \sim \mathrm{Li}(x) := \int_2^x \frac{dy}{\ln(y)}.$$

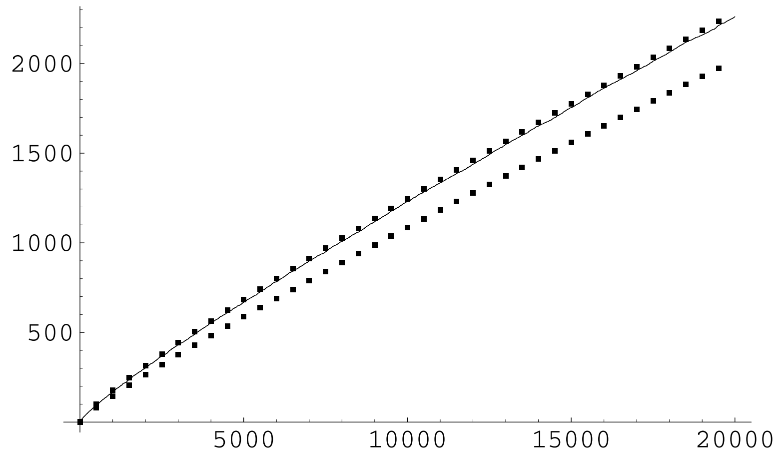

Figure 1: The number $\pi(x)$ of primes not exceeding x. For comparison: logarithmic integral (above) and function $x/\ln(x)$ (below).

See [42] for a short proof based on explicit formulae. In turn

$$\mathrm{Li}(x) \sim \frac{x}{\ln(x)},$$

but with slower convergence. Loosely speaking, one may thus say that near x, the primes have density $1/\ln(x)$.

Gauss discovered this law in 1792, at the age of fifteen, by studying prime number tables. Later on he found further evidence by listing the prime numbers up to *three million*.

In fact it is not hard to see that *if* the limit

$$\lim_{x\to\infty} \frac{\pi(x)}{\mathrm{Li}(x)} \tag{2}$$

exists, it must equal one. But it took over 100 years until in 1896 Hadamard and de la Valée Poussin proved independently convergence of (2) and thus the Prime Number Theorem [13]. Their proof was based on the Riemann zeta function, defined by the Dirichlet series

$$\zeta(s) := \sum_{n=1}^{\infty} n^{-s} \qquad (\Re(s) > 1), \tag{3}$$

which is absolutely convergent in that half-plane.

One obtains the Euler product form

$$\zeta(s) = \prod_{p\in\mathbb{P}} \frac{1}{1-p^{-s}} \qquad (\Re(s) > 1), \tag{4}$$

which exhibits more clearly its connection with the primes, by expanding the geometric series in (4), and using uniqueness of the prime factorization (1).

(Ordinary) *Dirichlet series* have the form

$$\sum_{n=1}^{\infty} a(n) n^{-s}$$

for some arbitrary *arithmetic function* $a : \mathbb{N} \to \mathbb{C}$, see, e.g. [2] for an introduction. They converge, resp. converge absolutely, in open half-planes of the form

$$\Re(s) > \sigma_c \quad \text{resp.} \quad \Re(s) > \sigma_a,$$

and diverge, resp. diverge absolutely, in the complement of their closures.

As can be read off from (3), ζ has a pole at one. So both its abscissa of convergence σ_c and its abscissa of absolute convergence σ_c equal one. However, $s = 1$ is the only pole of the Riemann zeta function, analytically continued to $\mathbb{C}$. This will follow from the functional equation and the

Theorem. $s = 1$ is the only pole of ζ in the half-plane $\Re(s) > 0$.

Proof. For $s > 0$

$$\begin{aligned} (1 - 2^{1-s}) \cdot \zeta(s) &= \sum_{n=1}^{\infty} n^{-s} - 2 \sum_{n=1}^{\infty} (2n)^{-s} \\ &= \sum_{n=1}^{\infty} \left[(2n-1)^{-s} - (2n)^{-s} \right] = \sum_{n=1}^{\infty} \mathcal{O}(n^{-1-s}) \neq \infty \end{aligned}$$

and similarly

$$(1 - 3^{1-s}) \cdot \zeta(s) \neq \infty \qquad (s > 0),$$

so that these Dirichlet series have $\sigma_c = 0$. The only common zero of $1 - 2^{1-s}$ and $1 - 3^{1-s}$ is $s = 1$, which proves the claim. $\square$

The *functional equation* for the Riemann zeta function ζ takes a simple form if one introduces the *complete zeta function*

$$\zeta_{\mathbb{A}}(s) := \pi^{-s/2} \cdot \Gamma(s/2) \cdot \zeta(s). \tag{5}$$

Theorem.

$$\zeta_{\mathbb{A}}(1-s) \equiv \zeta_{\mathbb{A}}(s) \tag{6}$$

Proof. The *Mellin transform* of a function f is given by

$$\mathcal{M}f(s) := \int_0^{\infty} (f(t) - f(\infty)) t^{s-1} dt. \tag{7}$$

Setting
$$f(t) := e^{-nt}, \qquad \mathcal{M}f(s) = \Gamma(s) \cdot n^{-s} \qquad (n > 0),$$
since for $n = 1$ this is the definition of the Gamma function, whereas for general n this follows by substitution of nt.

So $f(t) := g(e^{-t})$ for a power series $g(z) := \sum_n a(n)z^n$ is transformed into
$$\mathcal{M}f(s) = \Gamma(s) \cdot \sum_n a(n)n^{-s}.$$
In particular we obtain the identity
$$\zeta_{\mathbb{A}}(2s) = \mathcal{M}f(s) \quad \text{with} \quad f(t) := \tfrac{1}{2}\theta(it),$$
where θ denotes the Jacobi theta series, defined on the upper half plane
$$\mathbb{H} \equiv \mathbb{H}_\infty := \{z \in \mathbb{C} \mid \Im(z) > 0\}$$
by
$$\theta(z) := \sum_{n\in\mathbb{Z}} \exp(i\pi n^2 z).$$
Now applying the *Poisson summation formula*
$$\sum_{y\in\mathbb{Z}} \mathcal{F}h(y) = \sum_{x\in\mathbb{Z}} h(x)$$
for the Fourier transform $\mathcal{F}h$ of a Schwarz function h to $h_t(x) := \exp(-\pi t x^2)$, one sees that
$$\theta(-1/z) = \sqrt{-iz}\,\theta(z) \qquad (z \in \mathbb{H}), \tag{8}$$
so that the functional equation follows. □

Eq. (6) is indeed a remarkable identity, since it relates half-planes of divergence to half-planes of convergence.

Remark. In addition we note several things here.

- The Mellin transform is important in number theory since it relates power series and Dirichlet series.
- Fourier transformation, through Poisson summation, leads to the kind of duality encoded in the functional equation.
- In addition to (8), the Jacobi theta function meets $\theta(z+2) = \theta(z)$. So it is a so-called *modular function* (see [3, 40, 50]) for the subgroup of $\mathrm{SL}(2,\mathbb{Z})$ generated by the elements $\left(\begin{smallmatrix}0 & -1\\ 1 & 0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1 & 2\\ 0 & 1\end{smallmatrix}\right)$.

Here $M := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{R})$ acts on $\mathbb{H}$ by *Möbius transformations*:

$$z \mapsto \hat{M}(z) := \frac{az+b}{cz+d}. \tag{9}$$

Thus we see how the *modular group* $\Gamma \equiv \mathrm{PSL}(2,\mathbb{Z}) := \mathrm{SL}(2,\mathbb{Z})/\{\pm 1\!\!1\}$ and the modular domain $\Gamma \setminus \mathbb{H}$ come into play.

Theorem. The Riemann zeta function is a meromorphic function on $\mathbb{C}$ with a single, simple pole at $s = 1$, the *trivial zeroes* $s = -2n$, $n \in \mathbb{N}$, all other zeroes being contained in the *critical strip* $0 \leq \Re(s) \leq 1$, and

$$\zeta(\bar{s}) = \zeta(1-s) = 0 \quad \text{if} \quad \zeta(s) = 0.$$

Proof. We first show that there are no zeroes $\zeta(s) = 0$ with $\Re(s) > 1$. Zeroes of ζ become poles of its inverse

$$\frac{1}{\zeta(s)} = \prod_{p\in\mathbb{P}} (1-p^{-s}) = \sum_{n=1}^{\infty} \mu(n) n^{-s} \tag{10}$$

with the Möbius function[1]

$$\mu\left(\prod p^{\alpha_p}\right) = \begin{cases} (-1)^{\sum \alpha_p} & ,\alpha_p \leq 1 \\ 0 & ,\text{otherwise} \end{cases}.$$

But the Dirichlet series (10) converges for $\Re(s) > 1$, since $|\mu| \leq 1$.

The remaining statements follow immediately by inspection of the functional equation (6), see [2, 13]. □

Remark. The analytical proofs of the Prime Number Theorem basically reduce to a proof that the non-trivial zeroes of ζ are located in the *interior* of the critical strip.

Riemann Hypothesis (RH).
The non-trivial zeroes are on the *critical line* $\Re(s) = \frac{1}{2}$.

Despite constant efforts this remains an open question. If true, it implies a very smooth distribution of the primes [28], that is, small fluctuations of π away from Li:

$$(\mathrm{RH}) \qquad \Longrightarrow \qquad \pi(x) = \mathrm{Li} + \mathcal{O}(\sqrt{x} \cdot \ln(x)).$$

So much for the regular properties of the primes mentioned by Zagier.

The Dirichlet series (10) of $1/\zeta$ converges in the half-plane $\Re(s) > \frac{1}{2}$ if and only if the Riemann Hypothesis holds true.

[1] The Möbius function has found important generalizations in combinatorics, see [1].

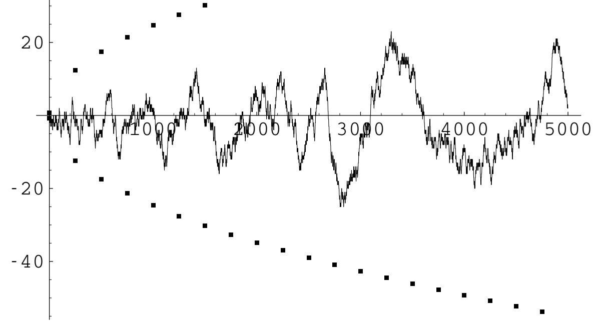

Figure 2: The sum function $M(x)$ of the Möbius μ function. For comparison: parabolae $\pm\sqrt{6\pi^{-2}x}$.

Thus setting $M(x) := \sum_{n\leq x}\mu(n)$, we obtain from partial summation

$$\frac{1}{\zeta(s)} = \int_1^\infty x^{-s}dM(x) = \int_1^\infty M(x)\cdot s\cdot x^{-s-1}dx$$

which converges for $\Re(s) > \sigma$ if $M(x) = \mathcal{O}(x^\sigma)$. In fact the (modified) *Mertens conjecture*

$$M(x) = \mathcal{O}\left(x^{1/2+\varepsilon}\right) \qquad (\varepsilon > 0) \tag{11}$$

is equivalent to the Riemann Hypothesis, see [53].

Now the graph of M looks very much like a random walk. This lead Denjoy (see [13]) and Good and Churchhouse in [18] to a probabilistic motivation for RH. Indeed (11) would follow with probability one if the $\mu(n)$ were i.i.d. random variables with the above distribution, since then $n \mapsto M(n)$ would correspond to a symmetric random walk.

However, arithmetical functions like μ are of course deterministic, and thus Good and Churchhouse remark that 'all our probability arguments are put forward in a purely heuristic spirit' [18].

One goal of these lectures is to show how to convert this analogy into a mathematical question concerning certain Markov processes. To speak metaphorically, it is clear that the primes are generated by a deterministic random number generator, but it is unclear how random this generator is.

Coming back to the functional equation (6), one may argue that $\zeta_{\mathbb{A}}$ is a function even more fundamental than ζ itself.

But why so? What is the significance of the Γ function appearing in its definition? This leads us to our next point.

3 Some p-adic Analysis

The Riemann zeta function is only the most prominent member of the family of so-called Dedekind zeta functions [57] which are associated to certain fields (algebraic extensions of $\mathbb{Q}$).

The relevant field for ζ is the quotient field $\mathbb{Q}$ of the ring $\mathbb{Z}$. We recall that the field $\mathbb{R}$ of real numbers is defined by forming Cauchy sequences of rationals, that is, by means which are not purely algebraic. Abstractly one starts from an *absolute value* $\|\cdot\| : \mathbb{Q} \to [0,\infty)$ with the properties

$$\|x\| = 0 \text{ iff } x = 0,$$

$$\|x_1 \cdot x_2\| = \|x_1\| \cdot \|x_2\| \quad \text{and} \quad \|x_1 + x_2\| = \|x_1\| + \|x_2\|,$$

takes the ring of Cauchy sequences w.r.t. to that absolute value and observes that their quotient w.r.t. the zero sequences forms an extension field of $\mathbb{Q}$.

In the case of the reals $\mathbb{Q}_\infty := \mathbb{R}$ one simply takes the absolute value $|\cdot|_\infty := |\cdot|$.

But this is not the only choice. Namely the p-adic field extensions $\mathbb{Q}_p$, for $p \in \mathbb{P}$, are constructed using the absolute value $|\cdot|_p$ with $|0|_p := 0$ and

$$|x|_p := p^{-\alpha_p} \quad \text{for} \quad x = \pm \prod_{q\in\mathbb{P}} q^{\alpha_p} \in \mathbb{Q}^* := \mathbb{Q} \setminus \{0\}.$$

We immediately notice the closure relation

$$\prod_{v\in\mathbb{P}\cup\{\infty\}} |x|_v = 1 \qquad (x \in \mathbb{Q}^*),$$

from which one gets the right impression that these are essentially the only absolute values (Theorem of Ostrowski, [55]).

Considering for a moment the ring $\mathbb{Z}_p \subset \mathbb{Q}_p$ of p-adic integers a, arising as limit points of Cauchy sequences $a^i \in \mathbb{Z}$ of ordinary integers, we notice that we can represent them as formal sequences

$$\sum_{n=m}^{\infty} a_n p^n \quad \text{with} \quad a_n \in \{0,\ldots,p-1\} \tag{12}$$

and $m = 0$, since the coefficients a_n^i eventually stabilizes. Elements of $\mathbb{Q}_p$ can still represented in the form (12), but now with arbitrary negative m.

From an algebraic point of view all these field extensions $\mathbb{Q}_v \supset \mathbb{Q}$, $v \in \mathbb{P}\cup\{\infty\}$, are equally important. They are all locally compact so that they have Haar measures dx_v, which are unique up to normalization. Thus

$$dx_\infty^* := \frac{dx_\infty}{|x|_\infty} \quad \text{resp.} \quad dx_p^* := \frac{p}{p-1}\frac{dx_p}{|x|_p} \qquad (p \in \mathbb{P})$$

are Haar measures on the multiplicative groups $\mathbb{Q}_p^*$.

Similar to the additive characters

$$\chi_{u,\infty} : \mathbb{Q}_\infty \to \mathbb{C}^* \qquad \chi_{u,\infty}(x) := \exp(2\pi i u x) \qquad (u \in \mathbb{Q}_\infty)$$

on $\mathbb{Q}_\infty = \mathbb{R}$, one has *additive characters* $\chi_{u,p}$ on $\mathbb{Q}_p$, given by

$$\chi_{u,p}(x) := \exp(2\pi i [ux]_p) \qquad (u \in \mathbb{Q}_p)$$

with the fractional part

$$[x]_p := \sum_{n=m}^{-1} a_n p^n \quad \text{for} \quad x = \sum_{n=m}^{\infty} a_n p^n \in \mathbb{Q}_p.$$

So we have formally Fourier Transformation

$$\mathcal{F}_v f(u) := \int_{\mathbb{Q}_v} f(x) \chi_{u,v}(x) dx_v$$

on this field and may ask ourselves about its eigenfunctions γ_v for eigenvalue 1. In the case of $\mathbb{R}$ we know that $\gamma_v(x) = \exp(-\pi x^2)$ is such an eigenfunction. For $p \in \mathbb{P}$

$$\gamma_p(x) := \begin{cases} 1, & |x|_p \leq 1 \\ 0, & |x|_p > 1 \end{cases}$$

solve the problem for the p-adic case, since for $|u|_p \leq 1$ the exponential equals 1, whereas for $|u|_p > 1$ we have effectively a sum over unit roots.

Similarly, for $s \in \mathbb{C}$

$$x \mapsto |x|_v^s$$

is a *multiplicative character* on $\mathbb{Q}_v$, $v \in \mathbb{P} \cup \{\infty\}$. Setting now

$$\zeta_v(s) := \int_{\mathbb{Q}_v} \gamma_v(x) |x|_v^s dx_v^* \qquad (v \in \mathbb{P} \cup \{\infty\}),$$

one calculates

$$\zeta_\infty(s) = \pi^{-s/2} \Gamma(s/2) \quad \text{and} \quad \zeta_p(s) = \frac{1}{1 - p^{-s}} \quad (p \in \mathbb{P}),$$

so that the complete zeta function has the form

$$\zeta_{\mathbb{A}} = \prod_{v \in \mathbb{P} \cup \{\infty\}} \zeta_v.$$

The index $\mathbb{A}$ indicates that $\zeta_{\mathbb{A}}$ is an *adelic* object, i.e. one in which all valuations of $\mathbb{Q}$ contribute, see [7, 30].

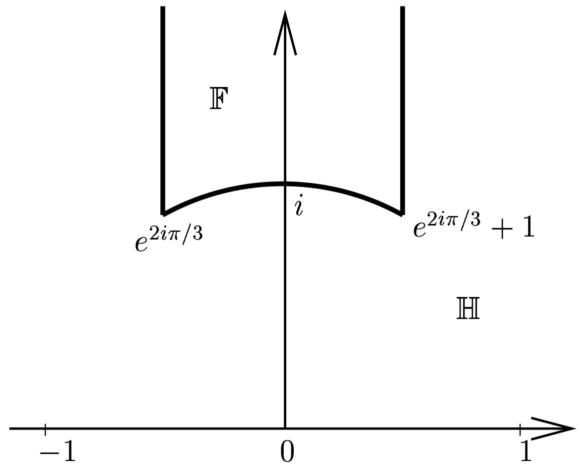

Figure 3: The modular domain $\mathbb{F}$

4 Scattering in the Modular Domain

Since the times of Minkowski [39], the geometrical aspects of number theory have turned out to be more and more important. We already saw the appearance of the group Γ, and we will now consider motion on the surface $\Gamma \setminus \mathbb{H}$.

The *modular domain* (see Fig. 3)

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

is a *fundamental domain* of Γ, that is, every $z \in \mathbb{H}$ can be carried by $M \in \Gamma$ 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

$$ds^2 = \frac{dx^2 + dy^2}{y^2}, \qquad (z \equiv x + iy \in \mathbb{H}) \tag{13}$$

of curvature -1, which is invariant under arbitrary Möbius transformations (9).

As can be seen from (13), the Laplace operator Δ on $\mathbb{H}$ has the form $\Delta = y^2(\partial_x^2 + \partial_y^2)$. The wave equation

$$\frac{\partial^2 u}{\partial t^2} = \left(\Delta + \tfrac{1}{4}\right) u \tag{14}$$

has the x–independent special solutions

$$h_\omega(z) \cdot e^{i\omega t} \quad \text{with} \quad h_\omega(z) := y^{\frac{1}{2}+i\omega} \qquad (z \in \mathbb{H}).$$

Starting from h_ω, we construct (Maass) *automorphic functions* $f : \mathbb{H} \to \mathbb{C}$, which are solutions of $-(\Delta + \frac{1}{4})f = \omega^2 f$ and are Γ–invariant ($f \circ \hat{M} = f$ for $M \in \Gamma$).

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 $\Gamma_\infty \backslash \Gamma$ and obtain the *Eisenstein series*

$$e(z,\omega) := \sum_{M \in \Gamma_\infty \backslash \Gamma} h_\omega(\hat{M}(z)) \tag{15}$$

which is known to converge for frequencies ω with $\Im(\omega) < -\frac{1}{2}$. By construction (15) is automorphic w.r.t. $z \in \mathbb{H}$. The *scattering matrix*[2] $\mathcal{S}(\omega)$ is defined with the help of the zeroth x-Fourier coefficient

$$e^{(0)}(y,\omega) := \int_{-\frac{1}{2}}^{\frac{1}{2}} e(x+iy,\omega)\,dx, \qquad y > 0,$$

namely

$$e^{(0)}(y,\omega) = y^{\frac{1}{2}+i\omega} + \mathcal{S}(\omega) y^{\frac{1}{2}-i\omega}. \tag{16}$$

Faddeev and Pavlov showed (see also [31]) that

Theorem [14].

$$\mathcal{S}(\omega) = \frac{\zeta_{\mathbb{A}}(2i\omega)}{\zeta_{\mathbb{A}}(1+2i\omega)}, \tag{17}$$

$\zeta_{\mathbb{A}}$ being the complete zeta function (5).

Proof. If $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$ is not in the coset space of the identity, then $c \neq 0$ and we can assume $c > 0$ because we are dealing with $\Gamma = \mathrm{SL}(2,\mathbb{Z})/\{\pm\mathbb{1}\}$. Then the other elements of the right coset of M are of the form $\left(\begin{smallmatrix} a-nc & b-nd \\ c & d \end{smallmatrix}\right)$, $n \in \mathbb{Z}$, so that there is a unique representative with $0 < a \leq c$. 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 $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$ with $0 < a \leq c$. Furthermore,

$$h_\omega(\hat{M}(z)) = \left(\Im \left(\frac{az+b}{cz+d} \right) \right)^{\frac{1}{2}+i\omega} = \frac{y^{\frac{1}{2}+i\omega}}{[(cx+d)^2 + c^2y^2]^{\frac{1}{2}+i\omega}},$$

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

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

[2] One should rather call it scattering *amplitude*. However, a similar construction for surfaces with n cusps leads to an $n \times n$ matrix [40].

and

$$
\begin{aligned}
e^{(0)}(y,\omega) &= y^{\frac{1}{2}+i\omega}\left(1+\sum_{c\in\mathbb{N}}\sum_{d\in\mathbb{Z},(c,d)=1}\int_{-\frac{1}{2}}^{\frac{1}{2}}[(cx+d)^2+c^2y^2]^{-\frac{1}{2}-i\omega}dx\right)\\
&= y^{\frac{1}{2}+i\omega}\left(1+\sum_{c\in\mathbb{N}}\sum_{\substack{1\le d\le c\\(c,d)=1}}\int_{-\infty}^{\infty}[(cx+d)^2+c^2y^2]^{-\frac{1}{2}-i\omega}dx\right)\\
&= y^{\frac{1}{2}+i\omega}+y^{\frac{1}{2}-i\omega}\frac{\Gamma(\frac{1}{2})\Gamma(i\omega)}{\Gamma(\frac{1}{2}+i\omega)}\sum_{c\in\mathbb{N}}\frac{\varphi(c)}{c^{1+2i\omega}} \qquad (18)
\end{aligned}
$$

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

By (16) this implies the Faddeev-Pavlov formula (17) for the scattering matrix $\mathcal{S}(\omega)$, since the quotient $\zeta(s-1)/\zeta(s)$ has the Dirichlet series

$$
\frac{\zeta(s-1)}{\zeta(s)} = \sum_{n\in\mathbb{N}}\frac{\varphi(n)}{n^s}. \qquad \Box \qquad (19)
$$

Remark. Berry, Keating and others noted similarities between the spectra of operators whose underlying dynamics is ergodic, and the non-trivial zeroes of the Riemann zeta function [5].

Here the poles of the scattering amplitude of Δ are related to these zeroes, and the geodesic flow on $\Gamma\setminus\mathbb{H}$ is mixing. This explains why such a connection exists.

The appearance of the adelic zeta function in (17) may come as a surprise, since the modular surface $\Gamma\setminus\mathbb{H}$ is defined in terms of the upper half plane $\mathbb{H}\equiv\mathbb{H}_\infty$ which is the homogeneous space

$$
\mathbb{H} = \mathrm{SL}(2,\mathbb{R})/\mathrm{SO}(2,\mathbb{R})
$$

(since $\mathrm{SO}(2,\mathbb{R})\subset\mathrm{SL}(2,\mathbb{R})$ is the isotropy group of $\sqrt{-1}\in\mathbb{H}$). So $\mathbb{H}_\infty$ is a *real* object.

However $\Gamma\setminus\mathbb{H}$ may be defined in purely adelic terms. The p-adic analogue $\mathbb{H}_p = \mathrm{PGL}(2,\mathbb{Q}_p)/\mathrm{PGL}(2,\mathbb{Z}_p)$ of $\mathbb{H}_\infty$ is naturally identified with the vertex set of the $(p+1)$–regular tree, and the scattering matrix on the modular domain has the form

$$
\mathcal{S} = \prod_{v\in\mathbb{P}\cup\{\infty\}}\mathcal{S}_v,
$$

the $\mathcal{S}_v$ being the scattering matrices on $\mathbb{H}_v$, see [7].

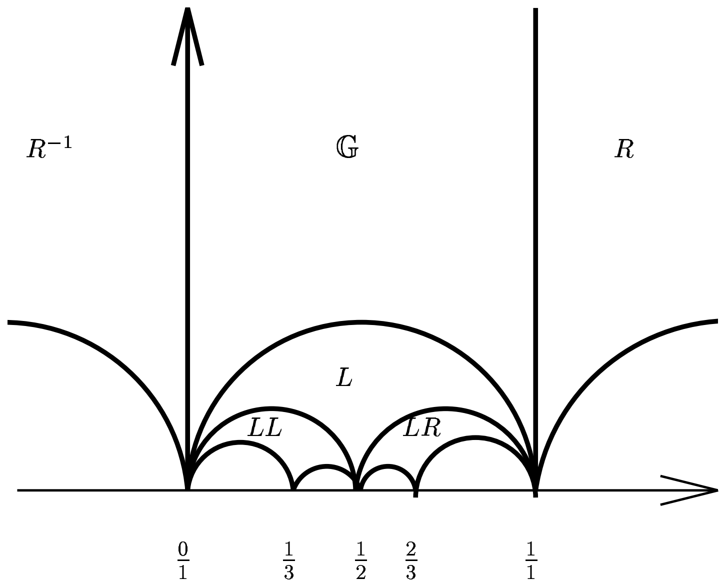

Figure 4: The fundamental domain $\mathbb{G}$ of the Farey tesselation

Remark. Motivated by [21] and [52], in [6] Bost and Connes considered the abstract statistical mechanical system whose partition function is the Riemann zeta function. A relation with the Riemann Hypothesis was stated in [8], see also [20].

It is unknown to me whether their (adelic) constructions are directly related to scattering theory and the approach presented here.

Now we will give a semiclassical geometrical interpretation of the Faddeev-Pavlov formula in terms of geodesics [26].

5 The Farey Tesselation

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 $S = \left(\begin{smallmatrix} 0 & 1 \\ -1 & 1 \end{smallmatrix}\right) \in \mathrm{SL}(2,\mathbb{Z})$, $\pm S \in \Gamma$ generating an order three subgroup.

Instead we will consider the *Farey tesselation* [51] of $\mathbb{H}$, whose fundamental domain

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

is the interior of a geodesic triangle, see Fig. 4.

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

The images of $\mathbb{G}$ w.r.t. the matrices in

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

form a tesselation of the strip $\{z \in \mathbb{H} \mid 0 \le \Re(z) \le 1\}$, since they map the cusps $(0, 1, \infty)$ of $\mathbb{G}$ onto themselves resp. onto

$$\left(\frac{b}{d}, \frac{a+b}{c+d}, \frac{a}{c}\right) \quad \text{with} \quad 0 \le \frac{b}{d} < \frac{a+b}{c+d} < \frac{a}{c} \le 1.$$

On the other hand the sum in the second line of (18) equals

$$\sum_{\hat{c} \in \mathbb{N}} \sum_{1 \le \hat{d} \le \hat{c},\, (\hat{c},\hat{d})=1} \hat{c}^{-1-2i\omega} = \sum_{\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}} (c+d)^{-1-2i\omega} \tag{20}$$

since there is a unique $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}$ with $c = \hat{c} - \hat{d}$ and $d = \hat{d}$.

To every $M = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathcal{U}$ we associate the *scattering geodesic* $c_{p/q}$ in $\mathbb{H}$ coming from ∞ and going to the middle of the three cusps of $\hat{M}(\mathbb{G})$. This geodesic is vertical, with real part p/q, where $p = a + b$ and $q = c + d$, see Fig. 5 a).

The first geodesic c_1 obtained in this way has real part one. We now compare the length of a geodesic $c_{p/q}$ with the the length of that geodesic. Although both lengths are, of course, infinite, the length differences turn out to be well-defined finite quantities.

So let in a general setting $\Phi^t : X \to X$ be the geodesic flow on the unit tangent bundle X of a Riemannian surface, with distance d. The (un)stable manifold of $x \in X$ is given by

$$W^{\pm}(x) := \left\{ y \in X \mid \lim_{t \to \pm\infty} d(\Phi^t(y), \Phi^t(x)) = 0 \right\}.$$

In the case of the geodesic flow for $\Gamma \backslash \mathbb{H}$ and x a point on the geodesic flow line corresponding to c_1 any scattering geodesics flow line in X (corresponding to a line in $\mathbb{H}$ with real part p/q) has unique points $y^{\pm} \in W^{\pm}(x)$.

Definition. The *time delay* $T(p/q)$ of the scattering geodesic $c_{p/q}$ is defined by $\Phi^{T(p/q)}(y^-) = y^+$.

This definition does not depend on the choice of the point x on the reference geodesic c_1.

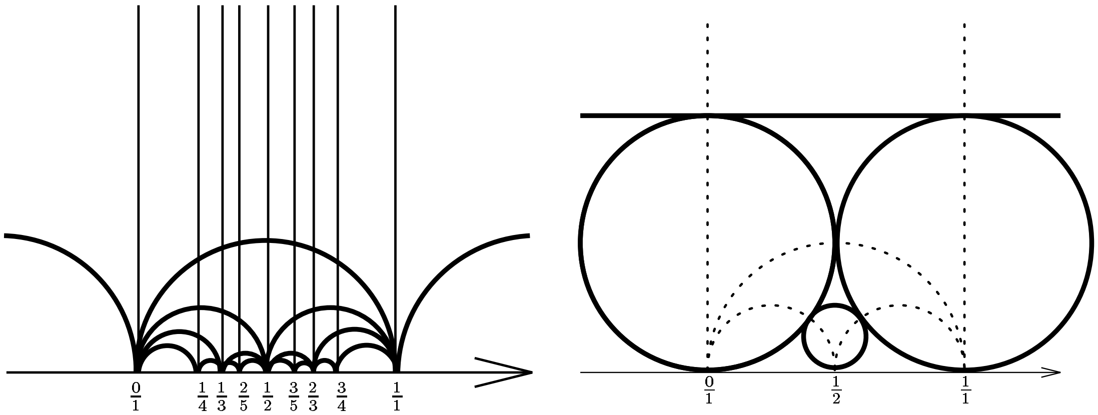

Figure 5: a) Geodesics related to the Farey tesselation. b) The Ford circles

Theorem. If $\gcd(p, q) = 1$, then the time delay equals $T(p/q) = 2\ln(q)$.

Proof. The *Ford circle* (see, e.g., [3]) in $\mathbb{H}$ with center at $\frac{p}{q} + \frac{i}{2q^2}$ and radius $\frac{1}{2}q^{-2}$ touches the cusp at $\frac{p}{q}$. The geodesics that cross that circle perpendicularly converge at the cusp. Thus it is the projection of the stable manifold of a point y^+ on the geodesic $c_{p/q}$, and y^+ projects to $p/q + i/q^2 \in \mathbb{H}$.

The Ford circle is the image of the horizontal line $\Im(z) = 1$ under the Möbius transformation $z \mapsto \hat{M}\hat{S}^{-1}(z)$, $M = \binom{a\,b}{c\,d}$. That horizontal line is the projection of the unstable manifold $W^-(y^-)$, where the point y^- lies on the geodesic $c_{p/q}$ and projects to $p/q + i$.

Equivalently, it is the image of the Ford circle $|z - (1 + i/2)| = \frac{1}{2}$ under the transformation $z \mapsto \hat{M}(z)$.

Its vertical metric distance to that line equals $2\ln(q)$, see Fig. 5 b). $\square$

Thus the term

$$(c + d)^{-2(\frac{1}{2}+i\omega)} = \exp\left(-T(p/q) \cdot (\tfrac{1}{2} + i\omega)\right)$$

in (20) may be interpreted as a phase shift of a partial wave due to the time delay of $T(p/q) = 2\ln(c + d)$ of the (unit speed) geodesic $c_{p/q}$ relative to the vertical geodesic c_1 in $\mathbb{H}$.

We thus have related the scattering amplitude $\mathcal{S}$ to a classical *dynamical zeta function*. Unlike conventional dynamical zeta functions here one does not sum over closed orbits but over scattering orbits.

Now comes a crucial point: These scattering geodesics connected to the Farey tesselation naturally come in families of 2^k members. One may see this by

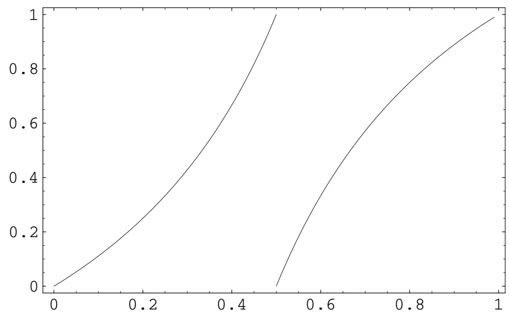

Figure 6: The map f.

observing that every matrix $M \in \mathcal{U} \setminus \{\mathbb{1}\}$ can be uniquely presented as

$$M = LX \quad \text{with} \quad L := \begin{pmatrix} 1\,0 \\ 1\,1 \end{pmatrix}, \, R := \begin{pmatrix} 1\,1 \\ 0\,1 \end{pmatrix} \quad \text{and} \quad X = \prod_i X_i$$

being a finite product of matrices $X_i = L$ resp. R (including the case $X = \mathbb{1}$).

Geometrically an application of $\hat{L}$ and $\hat{R}$ from the right means going to the neighbouring left resp. right Farey triangle which lies between the given one and the real axis, see Fig. 4.

Example. $\hat{L}\hat{L}\hat{R}\hat{L}(\mathbb{G})$ is the triangle with the cusps $(\frac{1}{3}, \frac{3}{8}, \frac{2}{5})$.

The middle cusps of $\hat{M}(\mathbb{G})$, $M \in \mathcal{U}$, are in bijection with the rationals $\mathbb{Q} \cap (0, 1]$. In order to recover the matrix X for a given $x \in \mathbb{Q} \cap (0, 1]$, we apply the map f, see Fig. 6,

$$f : [0, 1] \to [0, 1] \quad , \quad x \mapsto \begin{cases} \frac{x}{1-x} & 0 \le x \le \frac{1}{2} \\ 2 - \frac{1}{x} & \frac{1}{2} < x \le 1 \end{cases}$$

until $f^k(x) = \frac{1}{2}$, setting $X_i := L$ if $f^{i-1}(x) < \frac{1}{2}$ and $X_i := R$ if $f^{i-1}(x) > \frac{1}{2}$[3] .

Example. For $x := \frac{3}{8} < \frac{1}{2}$, $f(x) = \frac{3}{5} > \frac{1}{2}$, $f^2(x) = \frac{1}{3} < \frac{1}{2}$, $f^3(x) = \frac{1}{2}$, so that $X = LRL$.

The definition $f(x) = x/(1-x)$ of f for $0 \le x \le \frac{1}{2}$ coincides with the application of $\hat{L}^{-1}$ to that interval, whereas for $\frac{1}{2} < x \le 1$ we have $f(x) = 2 - 1/x = \hat{S}^{-1}\hat{R}^{-1}\hat{S}(x)$. (the conjugation by S is necessary in order to transform everything to the unit interval).

Iterated maps of the interval have been studied in depth. We may consider f as a continuously differentiable map on $\mathbb{R}/\mathbb{Z}$, and for $0 < x < 1$ it has the

[3] One may relate these preimages of 0 and the fixed points of f, using the Minkowski ?-function [39], whose inverse appears in the Appendix of [23].

expanding property $f'(x) > 1$. However at $x = 0$ the map is tangent to the diagonal, a fact which leads to particular properties, see [12, 15, 34].

Remark. f is also related to the *Gauss map*

$$g : [0,1] \to [0,1] \qquad x \mapsto 1/x - [1/x]$$

studied by Mayer [36, 37], Series [51], Lanford and Ruedin [29], and others in the context of *closed* geodesics in the modular domain. One such connection is given by the *Lewis functional equation* [33]

$$\lambda \cdot \psi(x) = \psi(x+1) + x^{-\beta}\psi(1+1/x),$$

since the free energy (28) of the number-theoretical spin chain (see below) equals $F(\beta) = -\ln(\lambda(\beta))/\beta$ [27].

At this point I would like to mention a relation between Farey fractions and the Riemann zeta function. The set F_n of *Farey fractions of order* n (see, e.g., [3]) is the set of (reduced) fractions $p/q \in (0,1]$, with $q \le n$ (so that $|F_n| = \sum_{k=1}^n \varphi(k)$ which is asymptotic to $3n^2/\pi^2$, see [2]). For example

$$F_5 = \left\{ \tfrac{1}{5}, \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{4}{5}, \tfrac{1}{1} \right\}.$$

For $q \in \mathbb{N}$ prime the relation

$$\sum_{\substack{0<p\le q:\\ \gcd(p,q)=1}} \exp\left(2\pi i \tfrac{p}{q}\right) = \mu(q) \qquad (q \in \mathbb{N}) \tag{21}$$

(μ being the Möbius function) follows, since the qth roots of unity have sum 0, and only the root 1 is missing. For arbitrary q one shows (21) by the inclusion-exclusion principle. From (21) we conclude

$$\sum_{x \in F_n} \exp(2\pi i x) = M(n) \equiv \sum_{q \le n} \mu(q).$$

Thus setting $N := |F(n)|$ and denoting by $x_l \in F_n$ the lth element in ascending order, by (11) the Riemann Hypothesis follows if

$$M(n) = \sum_{l=1}^{N} \left(\exp(2\pi i x_l) - \exp(2\pi i \tfrac{l}{N}) \right) = \mathcal{O}\left(n^{1/2+\varepsilon}\right),$$

thus in particular if the Farey fractions x_l of order n have a small mean distance

$$\sum_{l=1}^{N} \left| x_l - \tfrac{l}{N} \right| = \mathcal{O}\left(n^{1/2+\varepsilon}\right)$$

$k=0$	1																①
$k=1$	1								2								①
$k=2$	1				3				2				3				①
$k=3$	1		4		3		5		2		5		3		4		①
$k=4$	1	5	4	7	3	8	5	7	2	7	5	8	3	7	4	5	①

Figure 7: Pascal's triangle with memory.

to the points l/N and are thus quite equidistributed in $(0,1]$ (Franel and Landau, see [13]).

The $M \in \mathcal{U}$ with $M = \mathbb{1}$ or of the form

$$M = L \prod_{i=1}^{l} X_i$$

with $0 \le l < k$ lead to cusps $0 < x_1 < \ldots < x_{2^k} = 1$ which make up the set $\mathcal{F}_k$ of so-called *modified Farey fractions of order* k, with $\mathcal{F}_k \supset F_k$.

For example

$$\mathcal{F}_3 = \left\{ \tfrac{1}{4}, \tfrac{1}{3}, \tfrac{2}{5}, \tfrac{1}{2}, \tfrac{3}{5}, \tfrac{2}{3}, \tfrac{3}{4}, \tfrac{1}{1} \right\}.$$

We now concentrate on the denominators of $\mathcal{F}_k$. This leads us to classical statistical mechanics.

6 The Number-Theoretical Spin Chain

Since $|\mathcal{F}_k| = 2^k$, it is natural to enumerate the denominators by elements of the additive group

$$\mathbf{G}_k := (\mathbb{Z}/2\mathbb{Z})^k, \quad \text{with} \quad \mathbb{Z}/2\mathbb{Z} = (\{0,1\},+).$$

We then inductively set

$$\mathbf{h}_0 := 1, \quad \mathbf{h}_{k+1}(\sigma,0) := \mathbf{h}_k(\sigma) \quad \text{and} \quad \mathbf{h}_{k+1}(\sigma,1) := \mathbf{h}_k(\sigma) + \mathbf{h}_k(1-\sigma), \tag{22}$$

where $\sigma = (\sigma_1, \ldots, \sigma_k) \in \mathbf{G}_k$ and $1 - \sigma := (1-\sigma_1, \ldots, 1-\sigma_k)$ is the inverted configuration. Writing the $\mathbf{h}_k(\sigma)$ in the row number k using the lexicographic order of the $\sigma \in \{0,1\}^k$, we obtain what could be called *Pascal's triangle with memory*, see Fig. 7. Like in the usual Pascal triangle one writes the sum of

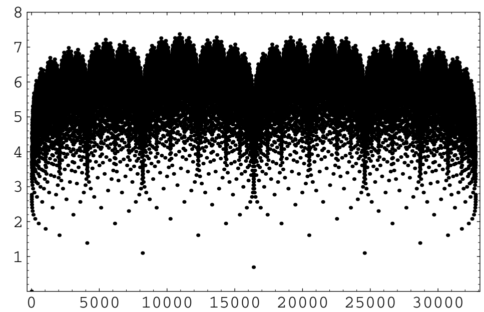

Figure 8: Graph of the energy function $\mathbf{H}_k$, $k = 15$, in lexicographic ordering of $\sigma \in \mathbf{G}_k$.

neighbouring integers in row no. k into the next row. But in addition one also copies the integers from row no. k to the $(k+1)$st row.

Notice that these sequences $\mathbf{h}_k(\sigma)$ of integers coincide with the denominators of the modified Farey sequence $\mathcal{F}_k$, except that 1 now has the zeroth instead the 2^kth position.

We now formally interpret $\sigma \in \mathbf{G}_k$ as a *configuration* of a spin chain with k spins and *energy function*

$$\mathbf{H}_k := \ln(\mathbf{h}_k),$$

see Fig. 8. Thus we may interpret

$$Z_k(s) := \sum_{\sigma \in \mathbf{G}_k} \exp(-s \cdot \mathbf{H}_k(\sigma))$$

as the *partition function* of that finite spin chain for *inverse temperature* s. The quotient

$$Z(s) := \frac{\zeta(s-1)}{\zeta(s)} \equiv \sum_{n=1}^{\infty} \varphi(n) n^{-s} \tag{23}$$

appearing in the scattering matrix (17) is simply the *thermodynamic limit*

$$\lim_{k \to \infty} Z_k(s) = Z(s) \qquad (\Re(s) > 2) \tag{24}$$

of partition functions $Z_k(s) = \sum_{n=1}^{\infty} \varphi_k(n) n^{-s}$, since $0 \leq \varphi_k(n) \leq \varphi(n)$ and $\varphi_k(n) = \varphi(n)$ if $n \leq k+1$. That *number-theoretical spin chain* was introduced in [23]. The *Gibbs measure* for inverse temperature $\beta \in \mathbb{R}$ assigns probabilities

$$\sigma \mapsto \frac{\exp(-\beta \mathbf{H}_k(\sigma))}{Z_k(\beta)} \qquad (\sigma \in \mathbf{G}_k) \tag{25}$$

to the configurations of the spin chain. We denote the expectation of a random variable by

$$\langle f \rangle_k (\beta) := \sum_{\sigma \in \mathbf{G}_k} f(\sigma) \frac{\exp(-\beta \mathbf{H}_k(\sigma))}{Z_k(\beta)} \qquad (f : \mathbf{G}_k \to \mathbb{R}).$$

To show that the analogy with statistical mechanics is not only formal, the *Fourier coefficients* of $\mathbf{H}_k$ were estimated in [23]. One notes that the dual group $\mathbf{G}_k^*$ of $\mathbf{G}_k$ is naturally isomorphic to $\mathbf{G}_k$, since the characters on $\mathbf{G}_k$ can be written in the form

$$\chi_t : \mathbf{G}_k \to \{-1, 1\} \quad , \quad \chi_t(\sigma) := (-1)^{\sum_{l=1}^k \sigma_i t_i} \qquad (t \in \mathbf{G}_k^*).$$

The Fourier coefficients

$$j_k(t) := -2^{-k} \sum_{\sigma \in \mathbf{G}_k} \mathbf{H}_k(\sigma) \cdot \chi_t(\sigma) \qquad (t \in \mathbf{G}_k^*)$$

of $-\mathbf{H}_k$ are called *interaction coefficients* in the statistical mechanics terminology, and

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

The negative mean $j_k(0)$ of $\mathbf{H}_k$ has special properties. In the thermodynamic limit it is asymptotic to $j_k(0) \sim -c \cdot k$ for some $c > 0$ [24], but it is the only coefficient whose value does not affect the Gibbs probability measure (25).

When we write $t \equiv (t_1, \ldots, t_k) \in \mathbf{G}_k^* \setminus \{0\}$ in the form

$$t = (0, \ldots, 0, 1, t_{l+1}, \ldots, t_{r-1}, 1, 0, \ldots, 0),$$

$s := r - l$ will be called the *size of* t, and $d := \min(l, k+1-r)$ its *distance* from the ends of the chain at 1 and k. Finally we say that t is *even* (*odd*) if $\sum_{i=1}^k t_i$ is *even* (*odd*). With these notations the following estimates were shown.

Theorem [23].

1. The even interactions decay exponentially in the size:

$$j_k(t) < 2^{-s} \qquad (t \in \mathbf{G}_k^* \setminus \{0\} \text{ even}), \tag{26}$$

whereas in the odd case one even has

$$j_k(t) < 2^{-(k-l)} \qquad (t \in \mathbf{G}_k^* \text{ odd}). \tag{27}$$

So odd interactions are small in comparison to the even ones except near the right end of the chain.

2. The interaction is *asymptotically translation invariant* in the sense that, up to a relative error which is exponentially small in the distance from the ends, the interactions only depend on the relative positions of the spins involved[4]:

$$0 \leq (j_{k+1}(0,t) - j_{k+1}(t,0)) \cdot 2^s < C \cdot 2^{-d} \qquad (t \in \mathbf{G}_k^*).$$

3. The interaction has a *thermodynamic limit* in the sense

$$0 \leq (j_{k+1}(0,t) - j_k(t)) \cdot 2^s < C \cdot 2^{-d} \qquad (t \in \mathbf{G}_k^* \setminus \{0\}).^5$$

For $\beta > 0$ the thermodynamic limit

$$F(\beta) := \lim_{k\to\infty} F_k(\beta) \quad \text{with} \quad F_k(\beta) := -\frac{1}{\beta \cdot k} \ln\left(Z_k(\beta)\right) \qquad (28)$$

of the *free energy* per spin exists [25].

4. The interaction is *ferromagnetic*, that is,

$$j_k(t) \geq 0 \qquad (t \in \mathbf{G}_k^* \setminus \{0\}). \qquad (29)$$

5. The *effective interaction*

$$A_k(l,r) := \sum_{t' \in \mathbf{G}_{r-l-1}^*} j_k(0,\ldots,0,1,t_{l+1},\ldots,t_{r-1},1,0,\ldots,0)$$

between spins at positions l and r decays quadratically with their distance $s = r - l$ in the bulk:

$$A_k(l,r) \leq \frac{1}{s^2} + \frac{2^{-(k-r)}}{s}. \qquad (30)$$

Remarks.

- Properties **1)–3)** are typical for many systems of statistical mechanics, and it should not be astonishing to find them in an application of the thermodynamic formalism.

  Since for $\Re(s) > 2$ the partition function Z of the infinite chain is finite, it is clear that $F(\beta) = 0$ in the low temperature range $\beta > 2$.

[4]For a related spin chain with an exactly translation invariant interaction see [22, 9].
[5]$C = 5$ covers both cases.

The free energy is important in statistical mechanics, since many physical quantities like the *internal energy* per spin

$$U(\beta) := \lim_{k\to\infty} U_k(\beta) \quad \text{with} \quad U_k(\beta) := \frac{1}{k}\left\langle \mathbf{H}_k \right\rangle_k(\beta) = \frac{d}{d\beta}(\beta F_k(\beta))$$

can be calculated from it. Points of non-analyticity of F are called *phase transitions*.

- The ferromagnetic property **4)**, however, comes as a surprise. There is no a priori reason why the Fourier transform of $-\mathbf{H}_k$ should be positive (except for the zeroth coefficient). I consider the ferromagnetic property as the main motivation for studying the ferromagnetic spin chain and return to that point in the next section.

- By (26) the individual coefficients decay exponentially w.r.t. the distance s of the spins involved in the interaction, but the number of such components increases exponentially, with the same rate.

  So **5)** sharpens (26) and (27) in the mean, and since that estimate is somewhat sharp, many-body interactions play an important role. In statistical mechanics it is known [45] that spin chains with a decay rate $s^{-\alpha}$ of the effective interaction show no phase transition if $\alpha > 2$. Since the number-theoretical spin chain has a phase transition [24], we are in the borderline situation.

Switching to a multiplicative representation $s_i(\sigma) := (-1)^{\sigma_i}$ of the ith spin, an important variable is the *mean magnetization* per spin

$$M_k(\beta) := \frac{1}{k}\sum_{i=1}^{k} \left\langle s_i \right\rangle_k(\beta)$$

and its thermodynamic limit $M(\beta)$. By analyzing a Perron-Frobenius operator with *PF* eigenvalue $\exp(-\beta \cdot F(\beta))$, the following statements were proved.

Theorem [10]. The only phase transition of the number-theoretical spin chain occurs for inverse temperature $\beta_{\mathrm{cr}} := 2$. For lower temperatures

$$F(\beta) = U(\beta) = 0 \quad \text{and} \quad M(\beta) = 1 \qquad (\beta > \beta_{\mathrm{cr}}),$$

whereas for high temperatures

$$U(\beta) \geq \tfrac{1}{4}(\beta - \beta_{\mathrm{cr}}) \quad (1 < \beta < \beta_{\mathrm{cr}}) \quad \text{and} \quad M(\beta) = 0 \quad (0 \leq \beta < \beta_{\mathrm{cr}}).$$

a)

```
111111111101100010001111011101100001000011111111011100000 0
000000000000000000000000000000001100000010010110000011111 1
001111101110000110100011111111111111111111111111111111111111
111111111111111111101111111111111011110111001111101000000011 0
111000000000000000000000000001001001100000000000000000000000 0
000000000011100011111010000011111011111111111111111110001101 11
111100000001110001111111111110000010010010111111001011111111 0
101110011111111110101101100000010101101111111111111101111111 1
111111111111111111111111111110000101111000100000000000000000 111
111110001111111110100000000111111101110011111110000000000000000 0
```

b)

```
110100010000001110000000100010011110100001011101000100011111 10
1011010011011111111111000111111111111101000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
000000000000000000000000000000000000000000000000000000000000
```

Figure 9: Typical configurations of $k = 600$ spins, for inverse temperature a) $\beta = 1.95$ and b) $\beta = 2.05$ near phase transition at $\beta_{\mathrm{cr}} := 2$.

Remarks.

1) Thus for low temperatures $\beta > \beta_{\mathrm{cr}}$, due to the long range of the interaction, the chain is in a frozen state, and with probability one only finitely many spins equal one, being located near the (left) end of the infinite chain.

The ferromagnetic property together with the GKS inequality (31) below abstractly implies that $M(\beta) \geq 0$ (if the thermodynamic limit exists). So the jump from $M(\beta) = 1$ to $M(\beta) = 0$ is the sharpest possible, see Fig. 9.

The phase transition is at most of second order, since the derivative of U is discontinuous at β_{cr}.

2) Amongst other things in [6] a, transition from a type I_∞ state for $\beta > 1$ to a type III factor for $\beta \leq 1$ was shown for the KMS_β states related to the Riemann zeta functions. In the language of C^* algebras, this corresponds to a similar phase transition as the one described above. Observe that in our case the partition function equals $\zeta(\beta - 1)/\zeta(\beta)$ so that the location of the phase transition is shifted by one.

7 Polymer Models and Ferromagnetism

Although there is no a priori reason for the positivity of the interaction, in a sense it did not come as a surprise. Instead, Kac (see his Comments in Pólya [44], pp. 424–426), Newman [41], Ruelle [47] and others had conjectured the existence of a ferromagnetic spin system related to the Riemann zeta function.

One motivation for that conjecture has been the *Lee-Yang circle theorem* of statistical mechanics. In its basic form it states that all zeroes of the partition function

$$Z(h) := \sum_{X \subset \Lambda} \exp(h|X|) \prod_{x \in X} \prod_{y \in \Lambda - X} a_{xy}$$

of a ferromagnetic ($a_{xy} = a_{yx} \in [-1, 1]$) Ising model occur at imaginary values of the external magnetic field h (see [48] for a proof).

Here we are interested in the zeroes of a partition function $Z(s)$ in the complex s plane, s being the inverse temperature. There exist variants of the Lee-Yang circle theorem which predict zero-free half-planes of the inverse temperature for certain ferromagnets, see, e.g., Ruelle [46].

Unfortunately, these theorems do not apply to our situation since our spin chain includes many-body interactions. It is known that the Lee-Yang theorem does not hold for general ferromagnetic systems with many-body interactions. Here, however, the interaction coefficients are of a rather special nature, as the proof of ferromagnetism (see below) will elucidate.

The proof of a first version of the circle theorem by Kac had been based on a technique applied by Pólya in [44]. There he took the asymptotics of the Fourier transform of ξ (with $\xi(iz) = \frac{1}{2}(z^2 - 1/4)\zeta_{\mathbb{A}}(\frac{1}{2} + z)$), and proved the Riemann Hypothesis for the inverse Fourier transform (which he called "verfälschte ξ-Funktion") of that function.

A second (admittedly vague) motivation to believe in the importance of the ferromagnetic property might be that positivity of a certain involution plays a central role in A. Weil's proof of the Riemann Hypothesis for curves over finite fields [42].

Thirdly there are *correlation inequalities* valid for expectations of ferromagnetic spin systems. As an example, in our additive notation the *GKS* inequalities [17] state that for $\beta \geq 0$ the expectations

$$\langle \chi_t \rangle_k (\beta) \geq 0 \qquad (t \in \mathbf{G}_k^*) \tag{31}$$

and

$$\langle \chi_{t_1} \chi_{t_2} \rangle_k (\beta) \geq \langle \chi_{t_1} \rangle_k (\beta) \cdot \langle \chi_{t_2} \rangle_k (\beta) \qquad (t_1, t_2 \in \mathbf{G}_k^*).$$

Thus much stronger information is available for ferromagnetic than for general spin systems.

After a first proof in [23] of the ferromagnetic property, a second, more conceptional proof in [19] was based on the polymer model technique. I shortly introduce polymer models and sketch the proof of ferromagnetism.

In an abstract setting one starts with a finite (or denumerable) set P whose elements are called *polymers*. Two given polymers $\gamma_1, \gamma_2 \in P$ may or may not overlap (be *incompatible*). Incompatibility is assumed to be a reflexive and symmetric relation on P.

Thus one may associate to a *k-polymer* $X := (\gamma_1, \ldots, \gamma_k) \in P^k$ an undirected graph $G(X) = (V(X), E(X))$ with vertex set $V(X) := \{1, \ldots, k\}$, vertices $i \neq j$ being connected by the edge $\{\gamma_i, \gamma_j\} \in E(X)$ if γ_i and γ_j are incompatible. Accordingly the k-polymer X is called *connected* if $G(X)$ is path-connected and *disconnected* if it has no edges ($E(X) = \emptyset$).

The corresponding subsets of P^k are called C^k resp. D^k, with $D^0 := P^0 := \{\emptyset\}$ consisting of a single element. Moreover $P^\infty := \bigcup_{k=0}^\infty P^k$ with the subsets $D^\infty := \bigcup_{k=0}^\infty D^k$ and $C^\infty := \bigcup_{k=1}^\infty C^k$. We write $|X| := k$ if $X \in P^k$.

Statistical weights or activities $z : P \to \mathbb{C}$ of the polymers are multiplied to give the activities $z^X := \prod_{i=1}^k z(\gamma_i)$ of k-polymers.

Definition. A system of statistical mechanics is called *polymer model* if its partition function Z has the form

$$Z = \sum_{X \in D^\infty} \frac{z^X}{|X|!}. \tag{32}$$

Then, up to the factor $-1/(\beta|\Lambda|)$, the free energy is given by

Theorem [16].

$$\ln(Z) = \sum_{X \in C^\infty} \frac{n(X)}{|X|!} z^X, \tag{33}$$

with $n(X) := n_+(X) - n_-(X)$, $n_\pm(X)$ being the number of subgraphs of $G(X)$ connecting the vertices of $G(X)$ with an even resp. odd number of edges.

For example, $n_+ = 1$ and $n_- = 4$ if the graph G is a quadrangle.

Example. We consider the 2-dim. nearest neighbour *Ising model* with energy function $H_\Lambda : \{\pm 1\}^\Lambda \to \mathbb{R}$ over a finite region $\Lambda \subset \mathbb{Z}^2$ (say, $\Lambda := \{0, \ldots, L\} \times \{0, \ldots, L\} \subset \mathbb{Z} \times \mathbb{Z}$)

$$H_\Lambda(s) := -\tfrac{1}{2} \sum_{\substack{i,j \in \Lambda \\ |i-j|_1 = 1}} s_i s_j. \tag{34}$$

Then (neglecting boundary effects) in a *low-temperature expansion*

$$\tfrac{1}{2} \exp(-2\beta|\Lambda|) \cdot Z_\Lambda(\beta) = \tfrac{1}{2} \sum_{s \in \{\pm 1\}^\Lambda} \exp\left(-\frac{\beta}{2} \sum_{i,j \in \Lambda, |i-j|_1 = 1} (s_i - s_j)^2/2 \right)$$

with partition function $Z_\Lambda(\beta) = \sum_s \exp(-\beta H_\Lambda(s))$ can be written in the form (32) where the polymers are the loops encircling regions of positive spin[6]. The activity equals

$$z(\gamma) := \exp(-2\beta|\gamma|), \tag{35}$$

where $|\gamma|$ is the length (number of edges) of γ.

In fact the low-temperature expansion is the first step in one method of *solving* (finding an explicit expression for the free energy) the 2-dim. Ising model.

There are several ways in which Dirichlet series can be related to polymer models [11]. However, a specific relation is useful to analyze the interaction of the ferromagnetic spin chain [19]:

Proof of the ferromagnetic property (29). In our case the energy function $\mathbf{H}_k = \log(\mathbf{h}_k)$ is the logarithm of a function (defined in (22)), whose Fourier coefficients

$$J_k(t) := 2^{-k} \sum_{\sigma \in \mathbf{G}_k} \mathbf{h}_k(\sigma) \cdot \chi_t(\sigma)$$

can be easily calculated: The set

$$P_k := \{p_{l,r}, p_l \in \mathbf{G}_k^* \mid 1 \le l, r \le k;\ l < r\}$$

of *polymers* in $\mathbf{G}_k^*$ is given by

$$p_{l,r}(i) := \delta_l(i) + \delta_r(i), \quad p_l(i) := \delta_l(i) \qquad (i = 1, \ldots, k). \tag{36}$$

We define their *support* by $\mathrm{supp}(p_{l,r}) := \{i \mid l \le i \le r\}$ for the *even* polymers and $\mathrm{supp}(p_l) := \{i \mid l \le i \le k\}$ for the *odd* ones. Two polymers $\gamma_1, \gamma_2 \in P_k$ are called overlapping or *incompatible* if $\mathrm{supp}(\gamma_1) \cap \mathrm{supp}(\gamma_2) \neq \emptyset$. So in particular all the odd polymers p_l, $1 \le l \le k$, are mutually incompatible.

Now an arbitrary group element $t = (t_1, \ldots, t_k) \in \mathbf{G}_k^*$ can be uniquely decomposed into the sum of $n := [\frac{1}{2}(|t| + 1)]$ compatible polymers:

$$t = \begin{cases} \sum_{i=1}^{n} p_{l_i,r_i} & , |t| \text{ even} \\ \sum_{i=1}^{n-1} p_{l_i,r_i} + p_{l_n} & , |t| \text{ odd.} \end{cases} \tag{37}$$

whose indices l_i, r_i may be assumed to be increasing in i and are determined by $t_{l_i} = t_{r_i} = 1$. As an example $(0,1,0,1,1,0) = p_{2,4} + p_5 \in \mathbf{G}_6^*$.

We write this decomposition in the short form $t = \sum_{i=1}^{n} \gamma_i$ and set the activity of a polymer $\gamma \in \mathbf{G}_k^*$ equal to $z(\gamma) := -3^{-|\mathrm{supp}(\gamma)|}$.

[6]Technically, one takes all loops which have no self-crossing and when meeting a vertex $v \in \Lambda$ from all sides, do not cross the main diagonal through v. Then two such polymers $\gamma_1, \gamma_2 \in P$ are *incompatible* if the loops contain a common edge, they cross, or meet at a vertex crossing the main diagonal.

Then

$$J_k(t) = (3/2)^k \prod_{i=1}^{n} z(\gamma_i),$$

so that the interaction coefficients $j_k(t)$ can be written as

$$\begin{aligned}
j_k(t) &= -2^{-k} \sum_{\sigma \in \mathbf{G}_k} \mathbf{H}_k(\sigma)\chi_t(\sigma) \\
&= -2^{-k} \sum_{\sigma \in \mathbf{G}_k} \ln \left[\sum_{s \in \mathbf{G}_k^*} J_k(s) \cdot \chi_\sigma(s) \right] \chi_t(\sigma) \\
&= -\delta_{t,0} \cdot k \ln(3/2) - 2^{-k} \sum_{\sigma \in \mathbf{G}_k} \ln \left[\sum_{X \in D_k^\infty} \frac{\tilde{z}_\sigma^X}{|X|!} \right] \cdot \chi_t(\sigma) \qquad (38)
\end{aligned}$$

where the redefined single-polymer activities $\tilde{z}_\sigma(\gamma)$, $\gamma \in P_k$ are given by

$$\tilde{z}_\sigma(\gamma) := z(\gamma) \cdot \chi_\sigma(\gamma) = -|z(\gamma)| \cdot (-1)^{\sigma \cdot \gamma}.$$

For all $\sigma \in \mathbf{G}_k$ the terms in the sum (38) are logarithms of expressions of the form (32), and we can use formula (33).

One notes in addition that the relevant graphs G are interval graphs[7], so that [19]

$$\mathrm{sign}(n(G)) = \begin{cases} 0 & G \text{ not connected} \\ -(-1)^{|V|} & G \text{ connected} \end{cases} .$$

Thus for $t \neq 0$ we only have positive contributions in (38), proving the ferromagnetic property. □

Note that not the free energy of the chain but its interaction coefficients were considered as abstract polymer models.

One may try to prove the Lee-Yang theorem for the number-theoretical spin chain, since its many-body interactions have that special combinatorial structure.

Most systems of statistical mechanics cannot be non-trivially presented as polymer models. Sometimes, however, a system can be presented in more than one way.

Example. In the *high temperature expansion* one uses the identity

$$e^{\pm\beta} = \cosh(\beta) \cdot (1 \pm \tanh(\beta))$$

[7] A graph $G = (V, E)$ is called *interval graph* if there exists an isomorphism $I : V \to \bar{V}$ between the set V of vertices and a set $\bar{V}$ of intervals in $\mathbb{R}$ such that $v_i, v_j \in V$ are connected by the edge $\{v_i, v_j\} \in E$ precisely if the corresponding intervals $I(v_i)$ and $I(v_j)$ have non-empty intersection.

in order to write the partition function Z_Λ of the 2-dim. nearest neighbour Ising model (34) in the form

$$Z_\Lambda(\beta) = (\cosh^2(\beta))^{|\Lambda|} \sum_{s\in\{\pm1\}^\Lambda} \prod_{i,k=0}^{L-1} (1 + x s_{i,k} \cdot s_{i,k+1}) \cdot (1 + x s_{i,k} \cdot s_{i+1,k})$$

with $x := \tanh(\beta)$. When the summation over s is performed, one obtains

$$Z_\Lambda(\beta) = (2\cosh^2(\beta))^{|\Lambda|} \cdot \sum_{G\in\Gamma} x^{|G|},$$

where Γ consists of the graphs G with vertex set Λ which have an even number of edges at every vertex, and $|G|$ denotes the number of edges.

Similar to the low-temperature expansion above, we can define the polymers to be the loop graphs γ, with activities

$$z(\gamma) := (\tanh\beta)^{|\gamma|}, \tag{39}$$

and represent the graphs $G \in \Gamma$ uniquely as disconnected multipolymers.

The *Kramers-Wannier duality* of the Ising model relates the partition function at temperatures β and

$$\beta^* := -\tfrac{1}{2}\ln(\tanh\beta).$$

The map $\beta \mapsto \beta^*$ is involutive, interchanges the activities (35) and (39), and thus leads to the functional equation

$$Z(\beta) = \tfrac{1}{2}(2\cosh(\beta)\sinh(\beta))^{|\Lambda|} Z(\beta^*).$$

Its fixed point is the point of phase transition.

In our language, it corresponds to an isomorphism of the polymer models for the low- and high temperature expansions. It is know since long [38] that, similar to (6), it can be understood on the basis of Poisson summation.

One open problem consists in an interpretation of the functional equation (6) of the Riemann zeta function in terms of a statistical mechanics duality.

8 Markov Chains and Ramanujan Graphs

The pole of ζ at $s = 2$ gives rise to a phase transition and leads to a divergence of the Dirichlet series (19) for smaller real parts. If we seek for a convergent Dirichlet series providing us information about ζ inside the critical strip, we must 'add weight with signs', instead of just 'adding weights', much as in the case of the Dirichlet series (10) for $1/\zeta$.

Instead of considering $1/Z$, let us change signs in the Euler product representation

$$Z(s) = \prod_{p\in\mathbb{P}} \frac{1-p^{-s}}{1-p^{1-s}} \qquad (\Re(s) > 2).$$

Then the "twisted" partition function

$$\tilde{Z}(s) := \prod_{p\in\mathbb{P}} \frac{1+p^{-s}}{1+p^{1-s}} = \frac{\zeta(s)\cdot\zeta(2(s-1))}{\zeta(s-1)\cdot\zeta(2s)} \tag{40}$$

has the Dirichlet series

$$\tilde{Z}(s) = \sum_{n=1}^{\infty} \lambda(n)\cdot\varphi(n)\cdot n^{-s} \qquad (\Re(s) > 2), \tag{41}$$

and the *Liouville function* $\lambda(\prod_{p\in\mathbb{P}} p^{\alpha_p}) := (-1)^{\sum_p \alpha_p}$ resembles the μ function in the Dirichlet series of $1/\zeta$. By (40) the pole of Z at $s=2$ is converted into a zero of $\tilde{Z}$, which in turn has a pole at $3/2$. The non-trivial zeroes s of ζ generate poles of $\tilde{Z}$ at $s+1$.

In [26] it was observed numerically that

$$\tilde{Z}_k(s) := \sum_{\sigma\in\mathbf{G}_k} \lambda(\mathbf{h}_k(\sigma))\cdot\exp(-s\cdot\mathbf{H}_k(\sigma))$$

approximates $\tilde{Z}(s)$ very well in the half-plane $\Re(s) > 3/2$, in fact much better than the truncations of the Dirichlet series (41).

In this half plane only the term $\zeta(s-1)$ in the product (40) has no convergent Dirichlet series. Thus a proof of convergence $\tilde{Z}_k \to \tilde{Z}$ for that half-plane would imply the Riemann Hypothesis.

Applying the heuristic random walk analogy of Sect. 2, we expect that in the thermodynamic limit the expectation

$$\langle\lambda\rangle_k = \mathcal{O}(2^{k/2+\varepsilon}) \qquad (\varepsilon > 0), \tag{42}$$

since the cardinality of our ensemble equals $|\mathbf{G}_k| = 2^k$. Here we use the notation

$$\langle f\rangle_k := \langle f\circ\mathbf{h}_k\rangle_k(0) = 2^{-k}\sum_{\sigma\in\mathbf{G}_k} f\circ\mathbf{h}_k(\sigma)$$

for expectations of arithmetic functions $f:\mathbb{N}\to\mathbb{R}$ on the ensemble of values $\mathbf{h}_k(\sigma)$, $\sigma\in\mathbf{G}_k$.

The conceptional advantage of the ensemble $\mathbf{h}_k(\sigma)$, $\sigma\in\mathbf{G}_k$, of integers over the ensemble $\{1,\ldots,k\}$ used in Sect. 2 lies in the relation with the ensemble $\mathbf{h}_{k+1}(\tau)$, $\tau\in\mathbf{G}_{k+1}$, imposed by the definition (22) of $\mathbf{h}_k$.

Setting for $m \in \mathbb{N}$ the function $\chi_m : \mathbb{N} \to \{0,1\}$ equal to $\chi_m(n) := 1$ if m divides n and zero otherwise, we have

$$\lambda = \prod_{p\in\mathbb{P}} \prod_{l=1}^{\infty} (-1)^{\chi_{p^l}} = \prod_{p\in\mathbb{P}} \prod_{l=1}^{\infty} (1 - 2\chi_{p^l}).$$

Furthermore for $m_1, m_2 \in \mathbb{N}$ one has

$$\chi_{m_1} \cdot \chi_{m_2} = \chi_m \qquad \text{with } m := \frac{m_1 m_2}{\gcd(m_1, m_2)}, \tag{43}$$

so that instead of estimating expectations (42), we may formally consider sums of expectations $\langle \chi_m \rangle_k$.

Now this leads us directly to the consideration of a *Markov chain* with state space

$$\Omega_m := (\mathbb{Z}/m\mathbb{Z}) \times (\mathbb{Z}/m\mathbb{Z}).$$

Namely χ_m descends to the function $\tilde{\chi} \in V_m$ in the Hilbert space

$$V_m := \{f : \Omega_m \to \mathbb{C}\} \quad \text{with inner product} \quad (f,g) := \sum_{\omega\in\Omega_m} \bar{f}(\omega) g(\omega),$$

$$\tilde{\chi}(a,b) := \delta_{a,0}, \qquad ((a,b) \in \Omega_m).$$

Setting the initial probability vector $v_{m,0} \in V_m$ of the Markov chain equal to

$$v_{m,0}(a,b) := \delta_{a,1} \cdot \delta_{b,1} \qquad ((a,b) \in \Omega_m),$$

and defining its transition matrix $T_m : V_m \to V_m$ by

$$T_M f := \tfrac{1}{2} \left(f \circ L^{-1} + f \circ R^{-1} \right) \qquad (f \in V_m),$$

with $L = \left(\begin{smallmatrix} 1 & 0 \\ 1 & 1 \end{smallmatrix}\right)$ and $R = \left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)$, we have

$$\langle \chi_m \rangle_k = (\tilde{\chi}_m, v_{m,k}). \tag{44}$$

for the vector

$$v_{m,k} := T_m v_{m,k-1} \qquad (k \in \mathbb{N}) \tag{45}$$

of probabilities at time k.

$v_{m,k}(a,b)$ can be interpreted as the normalized frequency of occurrence of the neighbouring pair (a,b) in the Pascal triangle with memory modulo m. For example, $v_{3,4}(0,2) = 1/16$, since the pair $(0,2)$ appears only once in the fourth line of Fig. 10.

$k=0$	1																1
$k=1$	1								2								1
$k=2$	1				0				2				0				1
$k=3$	1		1		0		2		2		2		0		1		1
$k=4$	1	2	1	1	0	2	2	1	2	1	2	2	0	1	1	2	1

Figure 10: Pascal's triangle with memory, modulo $m = 3$.

The thermodynamic limit $k \to \infty$ leads to the unique equilibrium state of the Markov chain, and the limit expectations are calculated as

$$\langle \chi_m \rangle_\infty := \lim_{k\to\infty} \langle \chi_m \rangle_k = \left(m \prod_{p\in\mathbb{P}, p|m} (1 + 1/p) \right)^{-1} . \tag{46}$$

We are interested in the *deviations* $\langle \chi_m \rangle_k - \langle \chi_m \rangle_\infty$ from the thermodynamic limit. If the $\chi_m \circ \mathbf{h}_k(\sigma)$, $\sigma \in \mathbf{G}_k$ were i.i.d. random variables, we would obtain $\langle \chi_m \rangle_k - \langle \chi_m \rangle_\infty = \mathcal{O}(2^{-k/2+\varepsilon})$. Instead, by using the relation

$$\langle \chi_m \rangle_k - \langle \chi_m \rangle_\infty = (\tilde{\chi}_m, T_m^k(v_{m,0}) - v_{m,\infty}), \tag{47}$$

we see that the exponential decay rate of that deviation equals the spectral radius of the transition matrix T_m (omitting the trivial *PF* eigenvalue 1). So one guesses that this spectral radius equals $1/\sqrt{2}$.

The actual spectral radius turns out to be larger, but the Markov chain $(\Omega_m, T_m, v_{m,0})$ carries information about the frequency of *pairs* $(a, b) \in \Omega_m$ of integers (mod m), whereas we are only interested in divisibility of a. So we do not use the full information encoded in the transition matrix T_n and may therefore reduce it.

Assuming for simplicity of the presentation that m is a prime, we observe that Ω_m is a two-dimensional vector space over the field $\mathbb{Z}/m\mathbb{Z}$, and that $\tilde{\chi}_m$ is the characteristic function of a line in that vector space. So we consider the reduced transition matrix $\bar{T}_m$ on the Hilbert space over the projective space $P^1(\mathbb{Z}/m\mathbb{Z})$ on which L and R act by Möbius transformations $z \mapsto z/(z+1)$ resp. $z \mapsto z+1$.

Example. In the basis corresponding to the representation $(1, \ldots, m, \infty)$ of $P^1(\mathbb{Z}/m\mathbb{Z})$ we have for $m = 3$,

$$L = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \quad , \quad R = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},$$

so that

$$\bar{T}_3 = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{48}$$

We could not show that the spectral radius of $\bar{T}_m - \bar{\Pi}_m$ (for the projector $\bar{\Pi}_m$ onto the Perron-Frobenius eigenvector $v_{m,\infty}$ of $\bar{T}_m$) equals $1/\sqrt{2}$. Instead, we proved the following

Theorem [27]. There is a $B < 1$ so that for all $m \in \mathbb{N}$

- All real eigenvalues $e \in \mathbb{R} \setminus \{1\}$ of $\bar{T}_m$ have modulus $|e| \leq B$.
- All non-real eigenvalues $e \in \mathbb{C} \setminus \mathbb{R}$ of $\bar{T}_m$ have modulus $|e| = 1/\sqrt{2}$.

In particular the spectral radius of $\bar{T}_m$ on the ortho-complement of the Perron-Frobenius eigenspace is bounded from above by B.

Remark. The important point here is the independence of the bound B from m. From (47) we see that the deviation of $\langle \chi_m \rangle_k$ from its thermodynamic limit (46) is of the form

$$\langle \chi_m \rangle_k - \langle \chi_m \rangle_\infty = \mathcal{O}(B^k). \tag{49}$$

This not only implies small fluctuations for the proportion of those $\sigma \in \mathbf{G}_k$ for which $\mathbf{h}_k(\sigma)$ is divisible by a given integer m.

In addition, by (43) and (46) these divisibility properties are only weakly correlated in the sense

$$\langle (\chi_{m_1} - \langle \chi_{m_1} \rangle_k) \cdot (\chi_{m_2} - \langle \chi_{m_2} \rangle_k) \rangle_k = \mathcal{O}(B^k)$$

for relatively prime integers m_1, m_2.

The proof of the theorem is based on the identity

$$(2\bar{T}_m + \bar{T}_m^{-1})^2 = 9B_m^2, \tag{50}$$

where B_m is a self-adjoint operator. (50) directly implies that the non-real spectrum is located on a circle of radius $1/\sqrt{2}$.

Further, if e is an eigenvalue of B_m, then $3(1-e)$ is an eigenvalue of the Laplacian of a certain three-regular graph G_m, which is constructed as follows.

The matrices $M_\pm \in \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$

$$M_+ := \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} = -L^{-1}R \quad \text{and} \quad M_- := \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix} = -LR^{-1} \tag{51}$$

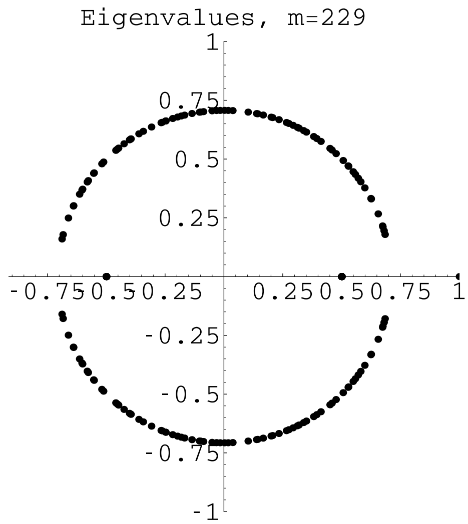

Figure 11: The spectrum of the reduced Markov transition matrix $\bar{T}_m$ for the 50th prime $m = 229$.

act by left transformations on the group $\mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$. For $m > 2$ the orbits of these actions are of size three, and the M_+-orbit and the M_--orbit through $g \in \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$ have only g in common.

Definition. We denote by V_+ (V_-) the set of M_+ (M_-)-orbits and consider

$$V := V_+ \cup V_-$$

as the vertex set of an undirected graph $G_m = (V, E)$. A pair $\{v_+, v_-\}$, $v_\pm \in V$ of vertices belongs to the set E of edges iff $v_+ \in V_+$, $v_- \in V_-$ and the orbits v_+ and v_- contain a common group element $g \in \mathrm{SL}(2, \mathbb{Z}/m\mathbb{Z})$.

Fig. 12 shows such a graph. A distinctive feature is its large girth (length of the shortest closed circuit).

The proof of the second statement in the above theorem is based on consideration of these graphs, using known properties of the so-called *Fell topology* on the set of (equivalence classes of irreducible unitary) representations of $\mathrm{SL}(2, \mathbb{Z})$, see [35].

Definition. A r-*regular* graph, i.e. one with r edges at every vertex, is called *Ramanujan* if the non-trivial spectrum $\sigma(\Delta) \setminus \{0, 2r\}$ of its Laplacian Δ is contained in the spectrum

$$\left[r - 2\sqrt{r-1}, r + 2\sqrt{r-1}\right]$$

of the r-regular tree.

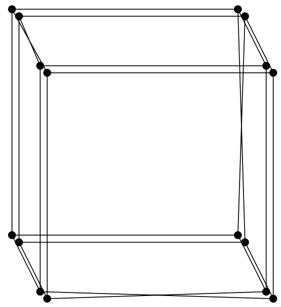

Figure 12: The graph G_3 for the group $\mathrm{SL}(2, \mathbb{Z}/3\mathbb{Z})$.

It may be shown that there is no family of finite r-regular graphs with diverging size meeting a better bound, so that Ramanujan graphs are in a sense optimal, see [35, 50, 54].

Theorem [27]. If the graph G_m has the Ramanujan property, then the spectral radius of $\bar{T}_m$ on the complement of the *PF* eigenspace equals $B = 1/\sqrt{2}$.

This would then indeed imply a decay of correlations (49) like in the case of an ensemble of i.i.d. random variables.

9 Final Remarks

We have seen how questions of number theory are related to geometrical and dynamical questions, and how these questions are reformulated using equilibrium statistical mechanics.

The general impression is that many different-looking concepts are in fact closely related. While this may be beautiful mathematics, one may ask whether it only leads to a vicious circle or it helps to answer old questions.

The most prominent such question is of course the one about the truth of the Riemann Hypothesis. Here one should be pessimistic, seen the long history of unsuccessful attempts. On the other hand, as a general rule, these attempts, although missing their direct aim, led to stimulating developments in different fields.

Applied to the present context, one is motivated to consider anew the Lee-Yang theorem, to ask about the significance of positive interaction function of a dynamical systems, cohomological questions of geodesic motion, to mention just a few aspects.

Moreover there are of course many simpler number-theoretical questions which one may try to answer using methods from dynamical systems and statistical mechanics.

References

[1] Aigner, M.: Combinatorial theory. New York : Springer 1979

[2] Apostol, T.M.: Introduction to Analytic Number Theory. Undergraduate Texts in Mathematics. New York: Springer 1976

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

[4] Baladi, V.: A Brief Introduction to Dynamical Zeta Functions. In: A. Knauf, Ya. Sinai: Classical Nonintegrability, Quantum Chaos. DMV Seminar, Vol. 27. Basel: Birkhäuser 1997

[5] Berry, M., Keating, J.P.: $H = xp$ and the Riemann Zeros. To appear in: Supersymmetry and Trace Formulae: Chaos and Disorder. J.P. Keating and I.V. Lerner, Eds. Plenum (1998)

[6] Bost, J.-B., Connes, A.: Hecke Algebras, Type III Factors and Phase Transitions with Spontaneous Symmetry Breaking in Number Theory. Selecta Mathematica, New Series, Vol. 1, No. 3, (1995), 411-457

[7] Brekke, L., Freund, P.: p-adic Numbers in Physics. Physics Reports **233**, 1–66 (1993)

[8] Connes, A.: Formule de trace en géométrie non-commutative et hypothèse de Riemann. C. R. Acad. Sci. Paris, Ser. I **323**, 1231–1236 (1996)

[9] Contucci, P., Kleban, P., Knauf, A.: A Fully Magnetizing Phase Transition. Max-Planck-Institute Preprint Nr.61/1998 (`http://www.mis.mpg.de/cgi-bin/preprints.pl`)

[10] Contucci, P., Knauf, A.: The Phase Transition of the Number-Theoretical Spin Chain. Forum Mathematicum **9**, 547–567 (1997)

[11] Contucci, P., Knauf, A.: The Low Activity Phase of Some Dirichlet Series. J. Math. Phys. **37**, 5458–5475 (1996).

[12] Cvitanović, P.: Circle Maps: Irrationally Winding. In: M. Waldschmidt, P. Moussa, J.M. Luck, C. Itzykson, Eds.: From Number Theory to Physics. Berlin, Heidelberg, New York: Springer 1992

[13] Edwards, H.M.: Riemann's Zeta Function. New York: Academic Press 1974

[14] Faddeev, 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)

[15] Feigenbaum, M., Procaccia, I., Tél, T.: Scaling properties of multifractals as an eigenvalue problem. Phys. Rev. A **39**, 5359–5372 (1989)

[16] Gallavotti, G., Martin-Löf, A., Miracle-Solé, S.: Some Problems Connected with the Description of Coexisting Phases at low Temperatures in the Ising Model. In: Statistical Mechanics and Mathematical Problems (Batelle, 1971) A. Lenard, ed., Lecture Notes in Physics **20**. Berlin: Springer 1973

[17] Ginibre, J.: Correlations in Ising Ferromagnets. In: Cargèse Lectures in Physics, Vol. 4. D. Kastler, ed., Gordon and Breach, London 1970

[18] Good, I.J., Churchhouse, R.F.: The Riemann Hypothesis and Pseudorandom Features of the Möbius Sequence. Mathematics of Computation **22**, 857–864 (1968)

[19] Guerra, F., Knauf, A.: Free Energy and Correlations of the Number-Theoretical Spin Chain. J. Math. Phys. **39**, 3188–3202 (1998)

[20] Harari, D., Leichtnam, E.: Extension du phénomène de brisure spontanée de symétrie de Bost-Connes au cas des corps globaux quelconques. Sel. Math., New Ser. **3**, 205–243 (1997)

[21] Julia, B.: Statistical Theory of Numbers. In: Number Theory and Physics. J.-M. Luck, P. Moussa, M. Waldschmidt, Eds., Proceedings in Physics **47**, Berlin: Springer 1988

[22] Kleban, P., Özlük, A.E.: A Farey Fraction Spin Chain. cond-mat/9808182 (1998). To appear in Commun. Math. Phys.

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

[24] Knauf, A.: Phases of the Number-Theoretical Spin Chain. Journal of Statistical Physics **73**, 423–431 (1993)

[25] Knauf, A.: On a Ferromagnetic Spin Chain. Part II: Thermodynamic Limit. J. Math. Phys. **35**, 228–236 (1994)

[26] Knauf, A.: Irregular Scattering, Number Theory, and Statistical Mechanics. In: Stochasticity and Quantum Chaos. Z. Haba et al, Eds. Dordrecht: Kluwer 1995

[27] Knauf, A.: The Number-Theoretical Spin Chain and the Riemann Zeroes. Commun. Math. Phys. **196**, 703–731 (1998)

[28] von Koch, H.: Ueber die Riemann'sche Primzahlfunction. Math. Ann. **55**, 441–464 (1902)

[29] Lanford, O.E., Ruedin, L.: Statistical Mechanical Methods and Continued Fractions. Helv. Phys. Acta **69**, 908–948 (1996)

[30] Lang, S.: Algebraic Number Theory. Graduate Texts in Mathematics **110**. New York: Springer 1986

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

[32] Lee, H., Levitov, L.: Universality in Phyllotaxis: A Mechanical Theory. In: Symmetry in Plants. D. Barabe, R. Jean, Eds., Singapore: World Scientific 1998

[33] Lewis, J.B., Zagier, D.: Period Functions and the Selberg Zeta Function for the Modular Group. Collection: The Mathematical Beauty of Physics, 83–97 Adv. Ser. Math. Phys. **24** (1996)

[34] Lopes, A.: The Zeta Function, Non-differentiability of Pressure, and the Critical Exponent of Transition. Adv. Math. **101**, 133–165 (1993)

[35] Lubotzky, A.: Discrete Groups, Expanding Graphs, and Invariant Measures. Progress in Mathematics **125**. Basel: Birkhäuser 1994

[36] Mayer, D.: On the Thermodynamic Formalism for the Gauss Map. Commun. Math. Phys. **130**, 311–333 (1990)

[37] Mayer, D.: The Thermodynamic Formalism Approach to Selberg's Zeta Function for $PSL(2,\mathbb{Z})$. Bull. AMS **25**, 55–60 (1991)

[38] McKean, H.P.: Kramers-Wannier Duality for the 2-Dimensional Ising Model as an Instance of Poisson's Summation Formula. J. Math. Phys. **5**, 775–776 (1964)

[39] Minkowski, H.: Zur Geometrie der Zahlen. Gesammelte Abhandlungen, Vol. 2. D. Hilbert, Ed.; Leipzig: Teubner 1911

[40] Müller, W.: Spektraltheorie automorpher Formen. In: Forschungsbericht 8, Akademie der Wissenschaften zu Berlin. Berlin, New York: de Gruyter 1993

[41] Newman, Ch.: Gaussian Correlation Inequalities for Ferromagnets. Z. Wahrscheinlichkeitstheorie verw. Gebiete **33**, 75–93 (1975)

[42] Patterson, S.J.: The Theory of the Riemann Zeta-function. Cambridge: Cambridge University Press 1995

[43] Pollicott, M.: Closed Geodesics and Zeta Functions. In: Ergodic Theory, Symbolic Dynamics and Hyperbolic Spaces. T. Bedford, M. Keane, C. Series, Eds. Oxford: Oxford University Press 1992

[44] Pólya, G.: Collected Papers, Vol. II: Locations of Zeros (R.P. Boas, ed.). Cambridge: M.I.T. Press 1974

[45] Ruelle, D.: Statistical Mechanics of a One-Dimensional Lattice Gas. Commun. Math. Phys. **9**, 267–278 (1968)

[46] Ruelle, D.: Some Remarks on the Location of Zeroes of the Partition Function for Lattice Systems. Commun. Math. Phys. **31**, 265–277 (1973)

[47] Ruelle, D.: Is Our Mathematics Natural? The Case of Equilibrium Statistical Mechanics. Bull. AMS **19**, 259–268 (1988)

[48] Ruelle, D.: Statistical Mechanics: Rigorous Results. Addison-Wesley, Redwood City 1989

[49] Ruelle, D.: Dynamical Zeta Functions: Where do they come from and what are they good for? In: Mathematical Physics X, K. Schmüdgen, Ed. Berlin: Springer 1991

[50] Sarnak, P.: Some Applications of Modular Forms. Cambridge, New York: Cambridge University Press 1990

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

[52] Spector, G.: Supersymmetry and the Möbius Inversion Function. Commun. Math. Phys. **127**, 239–252 (1990)

[53] Titchmarsh, E.C.: The Theory of the Riemann Zeta Function. London: Oxford University Press 1967

[54] Venkov, A., Nikitin, A.: Selberg Trace Formula, Ramanujan Graphs and some Problems of Mathematical Physics. Algebra Anal. **5**, 1–76 (1993)

[55] van der Waerden, B.: Algebra. Berlin: Springer 1993

[56] Zagier, D: The First 50 Million Prime Numbers. Math. Intell. **0**, 7–19 (1977)

[57] Zagier, D: Zetafunktionen und quadratische Körper. Berlin: Springer 1981