The Number-Theoretical Spin Chain and the Riemann Zeroes

Andreas Knauf*

June 1997

Abstract

It is an empirical observation that the Riemann zeta function can be well approximated in its critical strip using the Number-Theoretical Spin Chain. A proof of this would imply the Riemann Hypothesis. Here we relate that question to the one of spectral radii of a family of Markov chains. This in turn leads to the question whether certain graphs are Ramanujan.

The general idea is to explain the pseudorandom features of certain number-theoretical functions by considering them as observables of a spin chain of statistical mechanics. In an Appendix we relate the free energy of that chain to the Lewis Equation of modular theory.

1 Introduction

The Euler product formula

$$\zeta(s) \equiv \sum_{n=1}^{\infty} n^{-s} = \prod_{p\,\mathrm{prime}} \frac{1}{1-p^{-s}} \qquad (\mathrm{Re}(s) > 1)$$

for the Riemann zeta function and partial integration imply that

$$\frac{1}{\zeta(s)} = \sum_{n=1}^{\infty} \mu(n) n^{-s} = s \int_1^{\infty} \frac{M(x)}{x^{s+1}} dx, \qquad (\mathrm{Re}(s) > 1) \tag{1}$$

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

where $\mu : \mathbb{N} \to \{-1, 0, 1\}$ denotes the Möbius function and

$$M(x) := \sum_{n \le x} \mu(n).$$

Thus a Mertens type estimate

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

for all $\varepsilon > 0$ would imply convergence of (1) in the half plane $\mathrm{Re}(s) > \frac{1}{2}$ and thus the Riemann Hypothesis (RH)[1].

The values 1 and -1 of the Möbius function have equal densities $3/\pi^2$. This lead Good and Churchhouse in [8] to a probabilistic motivation for RH. Indeed (2) 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.

In this spirit, it is conjectured on the basis of the probabilistic law of the iterated logarithm for i.i.d. random variables that (2) is wrong for $\varepsilon = 0$. Indeed the original Mertens conjecture $|M(x)| < \sqrt{x}$ for $x > 1$ is known[2] to be wrong.

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 without any claim that they are mathematical proofs' [8].

In this paper we convert the above idea into a mathematical framework (which uses the Liouville function λ instead of the Möbius function).

This approach is based on a statistical mechanics interpretation of the Riemann zeta function.

In [10] we interpreted the quotient

$$Z(s) := \zeta(s-1)/\zeta(s) \tag{3}$$

of Riemann zeta functions for $\mathrm{Re}(s) > 2$ as the partition function of an infinite spin chain at inverse temperature s. In that half plane Z has the Dirichlet series

$$Z(s) = \sum_{n \in \mathbb{N}} \frac{\varphi(n)}{n^s}$$

with the Euler totient function $\varphi(n) := |\{j \in \{1, \ldots, n\} \mid \gcd(j, n) = 1\}|$. The quotient Z has been shown in [10] to be the *thermodynamic limit*

$$\lim_{k \to \infty} Z_k(s) = Z(s) \qquad (\mathrm{Re}(s) > 2) \tag{4}$$

[1]Eq. (2) is indeed equivalent to RH, see Titchmarsh [25], 14.25.

[2]See the article [21] by Odlyzko and te Riele.

$$
\begin{array}{lccccccccccccccccc}
k=0 & 1 & & & & & & & & & & & & & & & & 1 \\
k=1 & 1 & & & & & & & & 2 & & & & & & & & 1 \\
k=2 & 1 & & & & 3 & & & & 2 & & & & 3 & & & & 1 \\
k=3 & 1 & & 4 & & 3 & & 5 & & 2 & & 5 & & 3 & & 4 & & 1 \\
k=4 & 1 & 5 & 4 & 7 & 3 & 8 & 5 & 7 & 2 & 7 & 5 & 8 & 3 & 7 & 4 & 5 & 1
\end{array}
$$

Figure 1: Pascal's triangle with memory.

of the partition functions

$$Z_k(s) := \sum_{\sigma \in \{0,1\}^k} \exp(-s \cdot \mathbf{H}_k(\sigma))$$

of spin chains with k spins. The *energy function* $\mathbf{H}_k := \ln(\mathbf{h}_k)$ of that spin chain is defined inductively by

$$\mathbf{h}_0 := 1, \quad \mathbf{h}_{k+1}(\sigma, 0) := \mathbf{h}_k(\sigma) \quad \text{and} \quad \mathbf{h}_{k+1}(\sigma, 1) := \mathbf{h}_k(\sigma) + \mathbf{h}_k(1-\sigma), \tag{5}$$

the spin configuration $\sigma = (\sigma_1, \ldots, \sigma_k)$ being an element of the additive group $\mathbf{G}_k := (\mathbb{Z}/2\mathbb{Z})^k$ and $1-\sigma := (1-\sigma_1, \ldots, 1-\sigma_k)$ being the configuration with all spins inverted.

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. 1. Like in the usual Pascal triangle one writes the sum of 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 of integers coincide with the denominators of the modified Farey sequence.

For $n \leq k+1$ the multiplicity $\varphi_k(n) := |\{\sigma \in \mathbf{G}_k \mid \mathbf{h}_k(\sigma) = n\}|$ of n equals $\varphi(n)$. This implies (4), since

$$Z_k(s) = \sum_{n \in \mathbb{N}} \varphi_k(n) n^{-s}.$$

As has been worked out in [10, 11, 12], in the articles [5, 6] with Contucci, and the one with Guerra [9], this *Number-Theoretical Spin Chain* has the properties of typical systems considered in statistical mechanics. It has exactly one phase transition, at $s = 2$.

But from the point of view of number theory the most important point seems to be its *ferromagnetic* property. That is, the Fourier coefficients $j_k(t)$,

$t \in \mathbf{G}_k^* \cong \mathbf{G}_k$ of $-\mathbf{H}_k$ (*interaction coefficients* in the statistical mechanics terminology), with

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

are *positive*: $j_k(t) \geq 0$ for $t \neq 0$. The exceptional negative coefficient, namely the mean $j_k(0)$ of $-\mathbf{H}_k$, does not affect the *Gibbs measure*

$$\sigma \mapsto \exp(-s\mathbf{H}_k(\sigma))/Z_k(s) \qquad (\sigma \in \mathbf{G}_k)$$

of the spin chain for inverse temperature $s \geq 0$.

The ferromagnetic property is of interest in the context of the Riemann Hypothesis, since the Lee-Yang Theorem of statistical mechanics shows that the partition function of a ferromagnetic Ising system has a zero-free half-plane.

In [13] we noted that numerically the functions

$$\tilde{Z}_k(s) := \sum_{n=1}^{\infty} \lambda(n) \cdot \varphi_k(n) \cdot n^{-s} = \sum_{\sigma \in \mathbf{G}_k} \lambda(\mathbf{h}_k(\sigma)) \cdot \exp(-s \cdot \mathbf{H}_k(\sigma)) \qquad (6)$$

with the Liouville function

$$\lambda : \mathbb{N} \to \{\pm 1\}, \qquad \lambda\left(\prod_{p_i \text{ prime}} p_i^{\alpha_i}\right) := (-1)^{\sum_i \alpha_i}$$

well approximate the function

$$\tilde{Z}(s) := \sum_{n=1}^{\infty} \lambda(n) \cdot \varphi(n) \cdot n^{-s} = \frac{Z(2s) \cdot Z(2s-1)}{Z(s)} \qquad (7)$$

not only in the half plane $\mathrm{Re}(s) > 2$ of absolute convergence but even for $\mathrm{Re}(s) > 3/2$ (with $\tilde{Z}(s)$ being defined by analytical continuation). Comparing these statistical–mechanics ensembles with the finite Dirichlet series with the same number of terms obtained by truncation of the Dirichlet series (7), the numerical convergence properties of the $\tilde{Z}_k(s)$ are much better.

Clearly a convergence proof for the half plane $\mathrm{Re}(s) > 3/2$ would imply RH, since by (3) the non-trivial zeroes $s \in \mathbb{C}$ of ζ give rise to poles of $\tilde{Z}$ at $s + 1$.

In this paper we develop a framework supporting this empirical observation. In Sect. 2 we introduce a family of Markov chains with $n^2 \times n^2$ transition matrices T_n which control the divisibility by n of the values of $\mathbf{h}_k$.

In Sect 3. we show that T_n can be restricted to a subspace on which it is irreducible-aperiodic. However, the spectral radius of this reduced matrix

T_n^n is in general larger than $1/\sqrt{2}$ (which would be the expected value based on RH).

Thus in Sect. 4 we introduce a reduced matrix $\bar{T}_n$ (of size about n). We prove in Sect. 5 that all non-real eigenvalues of this doubly stochastic irreducible-aperiodic matrix have modulus $1/\sqrt{2}$.

In the course of the analysis of its real eigenvalues we are lead in Sect. 6 to a class of three-regular graphs which we conjecture to have the Ramanujan property (that is, their nontrivial spectrum is conjectured to be a subset of the spectrum of the three-regular tree).

In the last section we shortly draw some easy consequences from the Perron-Frobenius Theorem. However, we do not proceed in the analysis here, since further progress hinges on a proof of the Ramanujan property for the above graphs (or similar information).

In the Appendix the free energy of the Number-Theoretical Spin Chain is related to the solutions of the *Lewis Equation*

$$\psi(z) = \psi(z+1) + z^{-2s}\psi(1+1/z).$$

The holomorphic solutions of that equation on $\mathbb{C}\setminus(-\infty, 0]$ with $|\psi(0^+)| < \infty$ are in bijection with the even Maass wave forms, see Lewis [14].

Our approach based on the Number-Theoretical Spin Chain partly resembles the one followed by A. Connes. He interprets $\zeta(s)$ (instead of $Z(s)$) as the partition function of a statistical mechanics system at inverse temperature s, see [4].

Notation. We denote by $\mathbb{P} \subset \mathbb{N}$ the set of primes. $|S|$ is the cardinality of the set S.

Acknowledgement. I am most grateful to John Lewis who explained me his functional equation and showed how to generalize its relation with free energy from negative integral to arbitrary inverse temperatures.

2 The Markov Chain Construction

In (2) one considers the sums of the ensembles $(\mu(1), \ldots, \mu(k))$ in the limit $k \to \infty$. Here we estimate sums like $\sum_{\sigma \in \mathbf{G}_k} \lambda(\mathbf{h}_k(\sigma)) = \tilde{Z}_k(0)$, in order to gain some understanding of the functions $\tilde{Z}_k(s)$.

Thus we consider arithmetical functions $f : \mathbb{N} \to \mathbb{C}$ like λ as random variables $f \circ \mathbf{h}_k : \mathbf{G}_k \to \mathbb{C}$ with expectations

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

w.r.t. the normalized counting measure on the group $\mathbf{G}_k$.

In order to estimate the expectation $\langle\lambda\rangle_k$ of the Liouville function λ, we analyze the divisibility properties of $\mathbf{h}_k$. Therefore we start by considering for $m \in \mathbb{N}$ the functions $\chi_m : \mathbb{N} \to \{0,1\}$ and $c_m : \mathbb{N} \to \{-1,1\}$ given by $\chi_m(n) := 1$ for $m|n$, $\chi_m(n) := 0$ for $m \nmid n$ and $c_m := (-1)^{\chi_m} = 1 - 2\chi_m$. These functions are the building blocks of the arithmetical functions $\lambda_p : \mathbb{N} \to \{-1,1\}$, $p \in \mathbb{P}$, which equal $+1$ if the power of p in the prime factorization of the argument is even, and -1 if it is odd:

$$\lambda_p(m) := (-1)^{\sum_{l=1}^{\infty} \chi_{p^l}(m)} = \prod_{l=1}^{\infty} c_{p^l}(m). \tag{8}$$

Clearly for any argument m the sum and the product in (8) are effectively finite.

The Liouville function may be written as the product

$$\lambda = \prod_{p \in \mathbb{P}} \lambda_p$$

of these function (where one needs only take into account those primes p which are smaller than the argument).

We want to calculate thermodynamic limits

$$\langle f\rangle_\infty := \lim_{k\to\infty} \langle f\rangle_k \tag{9}$$

for the functions χ_m, c_m, λ_p and, finally, λ. Moreover if these limits are proven to exist, we would like to estimate how fast they are approached. Since $|\mathbf{G}_k| = 2^k$, an estimate of the form

$$\langle f\rangle_k - \langle f\rangle_\infty = \mathcal{O}_\varepsilon\left(\left(2^k\right)^{-1/2+\varepsilon}\right) \tag{10}$$

for all $\varepsilon > 0$ would be similar to the one expected for the μ-function as discussed in the Introduction and fit with the Riemann Hypothesis.

Our basic idea is to consider the groups $\mathbf{G}_k$ as finite probability spaces with the normalized counting measure, so that arithmetical functions composed with $\mathbf{h}_k$ are random variables on these probability spaces.

However, in order to estimate the k-dependence of these quantities, we embed all the finite groups $\mathbf{G}_k$ in the countably infinite probability space

$$\Omega := \mathbb{N} \times \mathbb{N},$$

with the help of the maps

$$I_k : \mathbf{G}_k \to \Omega, \qquad I_k(\sigma) := \left(\mathbf{h}_k(\sigma), \mathbf{h}_k(1-\sigma)\right).$$

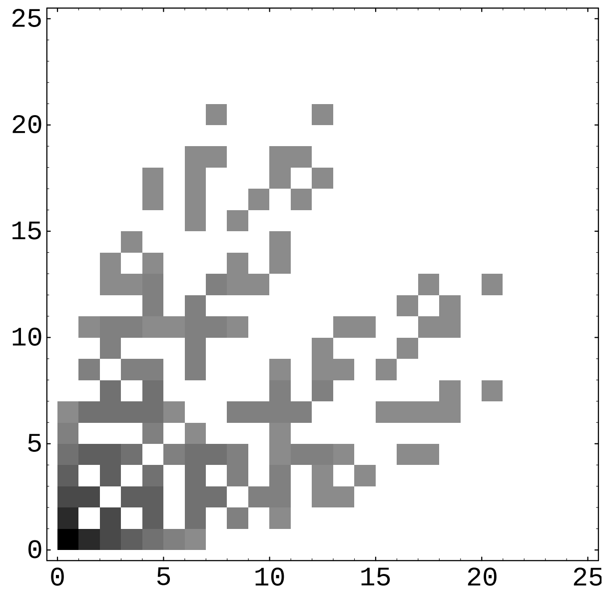

Figure 2: The images $I_k(\mathbf{G}_k) \subset \Omega$, for $k \leq 6$. Larger groups have lighter shading.

The I_k are indeed injective, have disjoint images $I_k(\mathbf{G}_k) \cap I_l(\mathbf{G}_l) = \emptyset$ for $k \neq l$, and

$$\bigcup_{k=0}^{\infty} I_k(\mathbf{G}_k) = \{(a,b) \in \Omega \mid \gcd(a,b) = 1\}$$

(see [10], Lemma 2.1 and 3.1). The images of the first groups are shown in Figure 2.

The image probability measures μ_k on Ω w.r.t. I_k give the elementary events the probabilities

$$\mu_k(\{(a,b)\}) = \begin{cases} 2^{-k} & (a,b) \in I_k(\mathbf{G}_k) \\ 0 & \text{otherwise} \end{cases} . \tag{11}$$

Now from the inductive definition (5) it follows that for all $k \in \mathbb{N}_0$, $\sigma \in \mathbf{G}_k$

$$(\mathbf{h}_{k+1}(\sigma,1), \mathbf{h}_{k+1}(1-\sigma,0)) = \tau^L(\mathbf{h}_k(\sigma), \mathbf{h}_k(1-\sigma))$$

and

$$(\mathbf{h}_{k+1}(\sigma,0), \mathbf{h}_{k+1}(1-\sigma,1)) = \tau^R(\mathbf{h}_k(\sigma), \mathbf{h}_k(1-\sigma))$$

with $\tau^L, \tau^R : \Omega \to \Omega$,

$$\tau^L(a,b) := (a+b,b) \text{ and } \tau^R(a,b) := (a,a+b)$$

denoting 'left' resp. 'right' addition.

Thus

$$\mu_{k+1} = \tfrac{1}{2}\left(\tau_*^L(\mu_k) + \tau_*^R(\mu_k)\right), \tag{12}$$

starting with the probability measure μ_0 concentrated on $(1,1) \in \Omega$.

If we denote by $\mathrm{pr} : \Omega \to \mathbb{N}$, $(a,b) \mapsto a$ the projection on the left factor, then $\mathrm{pr} \circ I_k = \mathbf{h}_k$ so that by (11)

$$\langle f \rangle_k = \sum_{\omega \in \Omega} f \circ \mathrm{pr}(\omega) \cdot \mu_k(\{\omega\}).$$

We want to estimate the expectations $\langle \chi_m \rangle_k$ or equivalently $\langle c_m \rangle_k$, since the Liouville function λ is a product of the c_m. However, in order to do this, one does not need to work on the infinite space Ω. Instead, for any $n \in \mathbb{N}$ with $m|n$ we can work on the space $\Omega_n := \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$. Then the functions $\chi_m \circ \mathrm{pr}$ on Ω are projectable w.r.t. the natural residue class map

$$X_n : \Omega \to \Omega_n, \qquad X_n((a,b)) := (a + n\mathbb{Z}, b + n\mathbb{Z}),$$

since they are constant on the preimages $X_n^{-1}(\omega)$, $\omega \in \Omega_n$.

We denote the projected functions by

$$\tilde{\chi}_m : \Omega_n \to \{0,1\} \quad \text{resp.} \quad \tilde{c}_m : \Omega_n \to \{-1,1\}.$$

These functions are thus elements of the n^2-dimensional Hilbert space

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

Moreover, the image measure $(X_n)_*(\mu_k)$ give probabilities

$$v_{n,k}(\omega) := \mu_k(X_n^{-1}(\{\omega\})) \tag{13}$$

to the elements $\omega \in \Omega_n$, and thus the expectation of a function $f \in V_n$ w.r.t. the image measure $(X_n)_*(\mu_k)$ equals the inner product $(f, v_{n,k})$.

In particular for integers m, n with $m|n$ one has

$$\langle \chi_m \rangle_k = (\tilde{\chi}_m, v_{n,k}). \tag{14}$$

3 The Matrices T_n

Since $X_n : \Omega \to \Omega_n$ acts $(\mathrm{mod}\ n)$, 'left' and 'right' addition τ^L, τ^R on Ω descend to maps $\tau_n^L, \tau_n^R : \Omega_n \to \Omega_n$

$$\tau_n^L(a,b) := (a+b, b), \quad \tau_n^R(a,b) := (a, a+b) \qquad (a,b) \in \Omega_n, \tag{15}$$

in the sense that $\tau_n^L \circ X_n = X_n \circ \tau^L$ and $\tau_n^R \circ X_n = X_n \circ \tau^R$.

But τ_n^L and τ_n^R are permutations. For a permutation $\tau \in \mathcal{S}(\Omega_n)$ we denote by $P_\tau : V_n \to V_n$, $P_\tau(f) := f \circ \tau^{-1}$ the permutation representation. Consider the endomorphism $T_n : V_n \to V_n$,

$$T_n := \tfrac{1}{2}(P_{\tau_n^L} + P_{\tau_n^R}).$$

The matrix representations of these endomorphisms w.r.t the orthonormal basis δ_x of characteristic functions of the points $x \in \Omega_n$ are denoted by the same symbol. As a convex combination of permutation matrices T_n is doubly stochastic, that is, its entries are nonnegative, and the sum of each column and each row equals one.

In fact, by (12) and by (13) T_n is the transition matrix of a finite *Markov chain* with state space Ω_n and probability vectors $v_{n,k}$, i.e.

$$T_n v_{n,k} = v_{n,k+1}, \qquad (k \in \mathbb{N}_0). \tag{16}$$

Our first goal in the spectral analysis of T_n is to find its ergodic sets $U_n^d \subset \Omega_n$. Thus we consider the orbits in Ω_n w.r.t. the action of the permutation subgroup generated by $\tau_n^L, \tau_n^R \in \mathcal{S}(\Omega_n)$.

Lemma 1 *For $d|n$ let*

$$\tilde{U}_n^d := \{(a,b) \in \Omega_n \mid n|ad \text{ and } n|bd\}.$$

Then $\tau_n^L(\tilde{U}_n^d) = \tau_n^R(\tilde{U}_n^d) = \tilde{U}_n^d$, and the cardinality $|\tilde{U}_n^d| = d^2$. For $d|n$ and $e|n$ one has

$$\tilde{U}_n^d \cap \tilde{U}_n^e = \tilde{U}_n^{\gcd(d,e)}.$$

Remarks 2

1. The property $n|ad$ for $a \in \mathbb{Z}/n\mathbb{Z}$ is independent of the chosen complete residue system so that the above definition is valid.
2. Obviously, $\tilde{U}_n^n = \Omega_n$ and $\tilde{U}_n^1 = \{(n,n)\}$. Furthermore, the map $\tilde{U}_n^d \to \tilde{U}_d^d$, $(a,b) \mapsto (ad/n, bd/n)$ is an isomorphism commuting with the τ^L and τ^R maps.

Proof. If $n|ad$ and $n|bd$ then $n|(a+b)d$, showing that $\tau_n^L(\tilde{U}_n^d) \subset \tilde{U}_n^d$ and $\tau_n^R(\tilde{U}_n^d) \subset \tilde{U}_n^d$. Equality holds since τ_n^L and τ_n^R are injective.

Furthermore, $\{a \in \mathbb{Z}/n\mathbb{Z} \mid n|ad\} = \{kn/d \mid k = 1, \ldots, d\}$ so that $|\tilde{U}_n^d| = d^2$.

If $d|n$ and $e|n$, then $n|cd$ and $n|ce$ implies $n|\gcd(cd, ce) = c\gcd(d,e)$, so that $\tilde{U}_n^d \cap \tilde{U}_n^e \subset \tilde{U}_n^{\gcd(d,e)}$. The converse inclusion is trivial. □

Lemma 3 *For $d|n$ let*

$$U_n^d := \tilde{U}_n^d - \bigcup_{p\in\mathbb{P},p|d} \tilde{U}_n^{d/p}.$$

Then

$$U_n^d = \{(a,b)\in\Omega_n \mid d = n/\gcd(a,b,n)\}, \qquad \tilde{U}_n^d = \dot{\bigcup}_{e|d} U_n^e, \tag{17}$$

$\tau_n^L(U_n^d) = \tau_n^R(U_n^d) = U_n^d$, *and the cardinality*

$$|U_n^d| = \sum_{e|d} \mu(d/e)e^2 = d^2 \cdot \prod_{p\in\mathbb{P},p|d} (1-p^{-2}). \tag{18}$$

Proof. We have

$$\tilde{U}_n^d = \{(a,b)\in\Omega_n \mid n|d\cdot\gcd(a,b)\} = \{(a,b)\in\Omega_n \mid n|d\cdot\gcd(a,b,n)\},$$

since $\gcd(n,\gcd(a,b)) = \gcd(a,b,n)$. By the definition of U_n^d, for $(a,b)\in U_n^d$ one has $n\nmid d'\gcd(a,b)$ for $d'<d$ and $d'|d$. This implies the equation $U_n^d = \{(a,b)\in\Omega_n \mid d = n/\gcd(a,b,n)\}$.

The same equation shows that $U_n^{d_1}\cap U_n^{d_2} = \emptyset$ if $d_1\neq d_2$, and that $\tilde{U}_n^d = \bigcup_{e|d} U_n^e$, proving the second part of (17).

The invariance of U_n^d w.r.t. the automorphisms τ_n^L and τ_n^R of the $\tilde{U}_n^d$ follows from

$$\begin{aligned}
\tau_n^L(U_n^d) &= \tau_n^L\left(\tilde{U}_n^d - \bigcup_{p\in\mathbb{P},p|d} \tilde{U}_n^{d/p}\right) = \tau_n^L\left(\tilde{U}_n^d\right) - \bigcup_{\substack{p\in\mathbb{P}\\ p|d}} \tau_n^L\left(\tilde{U}_n^{d/p}\right)\\
&= \tilde{U}_n^d - \bigcup_{\substack{p\in\mathbb{P}\\ p|d}} \tilde{U}_n^{d/p} = U_n^d
\end{aligned}$$

and similarly for τ_n^R.

The formula $|U_n^d| = \sum_{e|d}\mu(d/e)e^2$ for the cardinality follows from $|\tilde{U}_n^d| = \sum_{e|d}|U_n^e|$ and $|\tilde{U}_n^d| = d^2$, using the Möbius inversion formula (Theorem 2.9 of [1]).

The identity $\sum_{e|d}\mu(d/e)(e/d)^2 = \prod_{p\in\mathbb{P},p|d}(1-p^{-2})$ is certainly true for $d=1$ since then the product is empty. For $d>1$ let $p_1,\dots,p_r\in\mathbb{P}$ be the prime divisors of d. Then

$$\prod_{p\in\mathbb{P},p|d}(1-p^{-2}) = \prod_{i=1}^r (1-p_i^{-2}) = \sum_{k=0}^r (-1)^k \sum_{\substack{M\subset\{1,\dots,r\}\\ |M|=k}} \prod_{i\in M} p_i^{-2}$$

$$= \sum_{M\subset\{1,\ldots,r\}} \frac{\mu(\prod_{i\in M} p_i)}{(\prod_{i\in M} p_i)^2}$$
$$= \sum_{e'|d} \frac{\mu(e')}{(e')^2} = \frac{1}{d^2}\sum_{e|d}\mu(d/e)e^2,$$

since $\mu(e') = 0$ if e' contains a prime factor raised to a power ≥ 2. □

In particular we have obtained the decomposition

$$\Omega_n = \dot{\bigcup}_{e|n} U_n^e \tag{19}$$

of the state space Ω_n into disjoint τ_n^L- and τ_n^R-invariant subsets U_n^e. We will show now that these subsets cannot be further decomposed, i.e. that they are the orbits of the subgroup of permutations generated by τ_n^L and τ_n^R.

For $d|n$ let $V_n^d := \mathrm{span}(\{\delta_x \mid x \in U_n^d\}) \subset V_n$ be the subspace corresponding to the states in U_n^d and $\Xi_n^d : V_n \to V_n^d$ the orthogonal projector on that space.

By the invariance property $\tau_n^L(U_n^d) = \tau_n^R(U_n^d) = U_n^d$ of Lemma 3

$$T_n^d := T_n|_{V_n^d}$$

is an endomorphism of the subspace V_n^d and thus has a doubly stochastic matrix representation w.r.t. the basis vectors δ_x, $x \in U_n^d$. Again, we denote the matrix by the same symbol.

Lemma 4 *For $n \in \mathbb{N}$ and $d|n$ the matrix T_n^d is irreducible-aperiodic.*

Proof. By Remark 2) above it suffices to prove the lemma for the special case T_d^d, since T_n^d is mapped to T_d^d under the isomorphism $V_n^d \to V_d^d$, $\delta_{(a,b)} \mapsto \delta_{(ad/n,bd/n)}$.

We must show that for a certain $l \in \mathbb{N}$ the lth power of the matrix T_d^d has only strictly positive entries. But since T_d^d is the restriction $T_d|_{V_d^d}$ of the matrix $T_d = \frac{1}{2}(P_{\tau_d^L} + P_{\tau_d^R})$, one may equivalently show that for every pair (a_i, b_i), $(a_f, b_f) \in U_d^d$ of initial and final states there exists a chain $(a_k, b_k) \in U_d^d$, $k = 1, \ldots, l$ starting at $(a_1, b_1) := (a_i, b_i)$, ending at $(a_l, b_l) := (a_f, b_f)$ and following the rule

$$(a_{k+1}, b_{k+1}) := \tau_{I_k}(a_k, b_k), \qquad k = 1, \ldots, l-1$$

with some choice of indices $I_k \in \{L, R\}$ of the permutations.

Actually we can drop the condition that all these chains have a common length l, for the following reason. For all $a, b \in \mathbb{Z}/d\mathbb{Z}$ with $\gcd(a, d) =$

$\gcd(d,b) = 1$, the states (a,d) resp. (d,b) are in U_d^d. These states are fixed points of left addition τ_d^L resp. right addition τ_d^R. Therefore if we have shown that the matrix T_d^d is irreducible of equivalently that any given (a_i, b_i), $(a_f, b_f) \in U_d^d$ can be connected by *some* chain of states, we can connect (a_i, b_i) to, say, $(1,d)$, perform an arbitrary number of left additions and then connect $(1,d)$ to (a_f, b_f). This then implies the existence of chains between all states of U_d^d with a *common* length l.

Secondly, it is sufficient to show the existence of a chain from an arbitrary state $(a_i, b_i) \in U_d^d$ to $(1,1) \in U_d^d$, since then by the group property of the set of permutations generated by τ_d^L and τ_d^R there also exists a chain from $(1,1)$ to an arbitrary state $(a_f, b_f) \in U_d^d$.

We build the chain from (a_i, b_i) to $(1,1)$ by joining a chain from (a_i, b_i) to a state on the diagonal with a chain from that state to $(1,1)$.

To go from (a_i, b_i) to the diagonal, we employ the Euclidean algorithm: We set $(a_1, b_1) := (a_i, b_i)$ and assume without loss of generality that $a_1 > b_1$ (if $a_1 = b_1$ we are already on the diagonal, and if $a_1 < b_1$ we interchange a_k and b_k in the following construction). Setting $r_0 := a_1$, $r_1 := b_1$, by the Euclidean algorithm

$$\begin{aligned} r_0 &= q_1 r_1 + r_2 \\ r_1 &= q_2 r_2 + r_3 \\ &\dots \\ r_{n-2} &= q_{n-1} r_{n-1} + r_n \end{aligned}$$

with $0 < r_{i+1} < r_i$ for $i = 1, \dots, n-2$, $0 < q_i < d$ and $r_n = r_{n-1}$.

This implies for the states in U_d^d that

$$\begin{aligned} (r_0 + (d - q_1) r_1, r_1) &= (r_2, r_1) \\ (r_2, r_1 + (d - q_2) r_2) &= (r_2, r_3) \\ &\dots \end{aligned}$$

so that by first applying $d - q_1 > 0$ left additions, then $d - q_2 > 0$ right additions etc., we reach after finitely many (say, k_1) steps the element $(a_{k_1}, b_{k_1}) = (r_{n-1}, r_{n-1}) \in U_d^d$ on the diagonal.

Since $a_{k_1} = b_{k_1} = r_{n-1}$, we have $\gcd(a_{k_1}, b_{k_1}) = r_{n-1}$, and Lemma 3 implies $\gcd(a,b,d) = 1$ for $(a,b) \in U_d^d$, so that $\gcd(r_{n-1}, d) = 1$. This implies that r_{n-1}, considered as an element of $\mathbb{Z}/d\mathbb{Z}$, is invertible, i.e. $q r_{n-1} = 1 (\mathrm{mod}\ d)$ for some $q \geq 1$. Thus

$$(a_{k_1} + (q-1) b_{k_1}, b_{k_1}) = (1, b_{k_1})$$

so that $q-1$ left additions

$$(a_{i+1},b_{i+1}) := (a_i+b_i,b_i), \qquad i=k_1,\ldots,k_1+q-1,$$

lead to $(a_{k_2},b_{k_2})=(1,b_{k_1})$ with $k_2:=k_1+q-1$.

Finally, another $d+1-b_{k_1}$ right additions lead to $(a_{k_3},b_{k_3})=(1,1)$. $\square$

The Perron-Frobenius theorem, together with Lemma 4, implies the following facts:

1. The algebraic multiplicity of the eigenvalue 1 of the endomorphism T_n^d is one, and the vector
$$\mathbb{1}_n^d := \frac{1}{|U_n^d|}\sum_{x\in U_n^d}\delta_x \in V_n^d \tag{20}$$
is eigenvector w.r.t. that eigenvalue.

2. Denoting by Π_n^d the orthogonal projector on V_n^d to $\mathrm{span}(\mathbb{1}_n^d)$, the spectral radius
$$\mathrm{sr}\left(T_n^d-\Pi_n^d\right)<1. \tag{21}$$

By our direct sum decomposition,

$$T_n \cong \bigoplus_{d|n} T_n^d \quad \text{ on } V_n \cong \bigoplus_{d|n} V_n^d, \tag{22}$$

the algebraic multiplicity of the eigenvalue 1 of V_n being $\mathrm{d}(n)\equiv\sum_{d|n}1$, and the spectral radius

$$\mathrm{sr}(T_n-\sum_{d|n}\Pi_n^d\Xi_n^d)<1.$$

The smaller the spectral radius is, the faster the expectation $\langle f\rangle_k$ converges to its thermodynamical limit $\langle f\rangle_\infty$ as $k\to\infty$.

An inequality

$$\mathrm{sr}\left(T_n^d-\Pi_n^d\right)\le 1/\sqrt{2} \tag{23}$$

would be in accordance with (10) but is *not* valid in general (the first $n\in\mathbb{N}$ for which the spectral radius of T_n^n is strictly larger than $1/\sqrt{2}$ is $n=9$, the corresponding eigenvalues being roots of the polynomial $3+4x^2+16x^3+48x^4+64x^5+64x^6$).

The probability vectors $v_{n,k}\in V_n$ are actually elements of the subspace V_n^n, since $v_{n,0}$ is supported on $(1,1)\in U_n^n$, by the relation $v_{n,k}=(T_n)^k(v_{n,0})$ following from (16) and (22). So by (14)

$$\langle\chi_m\rangle_k = (\tilde{\chi}_m,(T_n^n)^k v_{n,0}). \tag{24}$$

But since by definition (20)

$$(\mathbb{1}_n^n, v_{n,k}) = (\mathbb{1}_n^n, (T_n)^k v_{n,0}) = ((T_n^*)^k \mathbb{1}_n^n, v_{n,0}) = (\mathbb{1}_n^n, v_{n,0}) = 1/|U_n^n|,$$

the Perron-Frobenius theorem implies

$$\lim_{k\to\infty} v_{n,k} = \mathbb{1}_n^n = \Pi_n^n v_{n,0}. \tag{25}$$

So the thermodynamic limit $\langle\chi_m\rangle_\infty$ defined in (9) is given by

$$\langle\chi_m\rangle_\infty = (\tilde{\chi}_m, \mathbb{1}_m^m), \tag{26}$$

and an explicit formula in terms of m will be given by evaluation of the r.h.s..

4 The Matrices $\bar{T}_n$

The violation of the estimate (23) seems to outlaw our probabilistic approach to the Riemann Hypothesis. However, since $\tilde{\chi}_m((a,b))$, with $(a,b) \in \Omega_n$, in (14) is independent of b, we do not use the full information encoded in the transition matrix T_n and may therefore reduce it.

In this section we thus analyze a new Markov chain with transition matrix $\bar{T}_n$ which is derived from T_n by lumping of states. $\bar{T}_n$ has smaller spectral radius than the old chain and suffices for our purpose.

As remarked in Section 2, we are not only interested in the thermodynamic limit (26), but mainly in the deviation $\langle\chi_m\rangle_k - \langle\chi_m\rangle_\infty$ from that limit for a spin chain of length $k < \infty$.

By (24) and (25) for all $m \in \mathbb{N}$ with $m|n$

$$\begin{aligned} \langle\chi_m\rangle_k - \langle\chi_m\rangle_\infty &= (\tilde{\chi}_m, (T_n^n)^k v_{n,0}) - (\tilde{\chi}_m, \Pi_n^n v_{n,0}) \\ &= (\tilde{\chi}_m, (T_n^n - \Pi_n^n)^k v_{n,0}), \end{aligned} \tag{27}$$

the second equation following from $T_n^n \Pi_n^n = \Pi_n^n T_n^n = \Pi_n^n$. From (27) we see that

$$|\langle\chi_m\rangle_k - \langle\chi_m\rangle_\infty| \le c \cdot (\mathrm{sr}(T_n^n - \Pi_n^n))^k$$

for some k-independent c, if T_n^n is semi-simple. By the Perron-Frobenius inequality (21) this deviation vanishes exponentially in k.

However we just remarked that for general n the estimate (23) does *not* hold true. On the other hand we may use the invariance property $\tilde{\chi}_m(xa, xb) = \tilde{\chi}_m(a,b)$ valid for the $x \in \mathbb{Z}/n\mathbb{Z}$ with $\gcd(x,n) = 1$ to improve the above estimate.

The ring $\mathbb{Z}/n\mathbb{Z}$ acts on the $\mathbb{Z}/n\mathbb{Z}$ module Ω_n by multiplication. If $n \notin \mathbb{P}$, then $\mathbb{Z}/n\mathbb{Z}$ is not a field (and Ω_n not a vector space over $\mathbb{Z}/n\mathbb{Z}$).

For general $n \in \mathbb{N}$ we consider the action

$$\alpha_n : U(\mathbb{Z}/n\mathbb{Z}) \times \Omega_n \to \Omega_n, \qquad (x,(a,b)) \mapsto (xa,xb)$$

of the multiplicative group

$$U(\mathbb{Z}/n\mathbb{Z}) := \{x \in \mathbb{Z}/n\mathbb{Z} \mid \gcd(x,n) = 1\}$$

which is of order $|U(\mathbb{Z}/n\mathbb{Z})| = \varphi(n)$.

That group action leaves the ergodic sets of T_n invariant:

Lemma 5 *For $n \in \mathbb{N}$, $d|n$ and $x \in U(\mathbb{Z}/n\mathbb{Z})$*

$$\alpha_n(x, U_n^d) = U_n^d.$$

Proof. We show first that

$$\alpha_n(x, \tilde{U}_n^d) = \tilde{U}_n^d, \qquad x \in U(\mathbb{Z}/n\mathbb{Z}).$$

Since $\tilde{U}_n^d$ is of the form

$$\tilde{U}_n^d = \{(k_1 n/d, k_2 n/d) \mid k_1, k_2 = 1, \ldots, d\},$$

it is obvious that $\alpha_n(x, \tilde{U}_n^d) \subset \tilde{U}_n^d$. But multiplication by $x \in U(\mathbb{Z}/n\mathbb{Z})$ is an automorphism of the finite set Ω_n so that the opposite inclusion holds, too.

By a argument similar to the one employed to prove τ_n^L, τ_n^R-invariance of U_n^d im Lemma 4, we conclude that $\alpha_n(x, U_n^d) = U_n^d$, too. $\square$

So we obtain a refinement of the partition (19) of Ω_n into the sets U_n^d by considering the orbits of the group action α_n.

We denote the restrictions of α_n to U_n^d by $\alpha_n^d : U(\mathbb{Z}/n\mathbb{Z}) \times U_n^d \to U_n^d$, the set of α_n-orbits by $\mathcal{O}_n := \Omega_n/U(\mathbb{Z}/n\mathbb{Z})$, and the set of α_n^d-orbits by $\mathcal{O}_n^d := U_n^d/U(\mathbb{Z}/n\mathbb{Z})$.

Lemma 6 *For $n \in \mathbb{N}$, $d|n$ and $\omega \in U_n^d$ the cardinality of the isotropy group of ω equals*

$$\left|\left\{x \in U(\mathbb{Z}/n\mathbb{Z}) \mid \alpha_n^d(x,\omega) = \omega\right\}\right| = \varphi(n)/\varphi(d), \tag{28}$$

and the cardinality of the orbit through ω equals

$$\left|\alpha_n^d(U(\mathbb{Z}/n\mathbb{Z}), \omega)\right| = \varphi(d). \tag{29}$$

The number of α_n^d-orbits in U_n^d equals

$$\left|\mathcal{O}_n^d\right| = d \prod_{p \in \mathbb{P}, p|d} (1 + 1/p). \tag{30}$$

Proof. The proof is based on the following fact (see, e.g., Thm 5.33 of [1]): Consider a reduced residue system $\hat{U}(\mathbb{Z}/n\mathbb{Z})$ for $U(\mathbb{Z}/n\mathbb{Z})$ and let $d|n$. Then $\hat{U}(\mathbb{Z}/n\mathbb{Z})$ is the disjoint union of $\varphi(d)$ sets, each of which consists of $\varphi(n)/\varphi(d)$ numbers congruent to each other (mod d).

To see how formula (28) for the cardinality of the isotropy group follows, we consider the elements $x \in \hat{U}(\mathbb{Z}/n\mathbb{Z})$ which are congruent to one (mod d). By the above fact, that set has cardinality $\varphi(n)/\varphi(d)$ and clearly is a subset of the isotropy group of ω, since $\omega \in U_n^d$ is of the form $(k_1 n/d, k_2 n/d)$.

On the other hand, let $x \in \hat{U}(\mathbb{Z}/n\mathbb{Z})$ be an element of the isotropy group of $\omega = (a, b) \in U_n^d$. Then $n|(x-1)a$ and $n|(x-1)b$ so that $n|(x-1)\gcd(a, b)$ and $n|(x-1)\gcd(a, b, n)$. But by (17) one has $\gcd(a, b, n) = n/d$ so that $n|(x-1)n/d$. This can only be the case if $x - 1$ is a multiple of d, which shows the opposite inclusion.

The fact cited above also implies formula (29) for the cardinality of the orbit through ω, since elements $x_1, x_2 \in \hat{U}(\mathbb{Z}/n\mathbb{Z})$ which are not congruent (mod d) lead to different points $\alpha_n^d(x_1, \omega) \neq \alpha_n^d(x_2, \omega)$ of the orbit.

Remembering the product representation $\varphi(d) = d\prod_{p\in\mathbb{P},p|d}(1 - 1/p)$ of the Euler totient, we obtain the formula (30) for the number of orbits in U_n^d by dividing the cardinality (18) of U_n^d through the (constant) cardinality (29) of the orbits:

$$\left|\mathcal{O}_n^d\right| = \frac{d^2 \cdot \prod_{p\in\mathbb{P},p|d}(1 - p^{-2})}{d \cdot \prod_{p\in\mathbb{P},p|d}(1 - p^{-1})} = d \cdot \prod_{p\in\mathbb{P},p|d} (1 + p^{-1}). \qquad \square$$

The isomorphism $U_n^d \to U_d^d$, $(a, b) \mapsto (ad/n, bd/n)$ maps α_n^d-orbits onto α_d^d-orbits and thus induces an isomorphism $\mathcal{O}_n^d \to \mathcal{O}_d^d$. So we need only consider the sets $\mathcal{O}_d^d$ of orbits.

- If $d \in \mathbb{P}$, then by (30) $|\mathcal{O}_d^d| = d + 1$, and $\mathcal{O}_d^d$ is isomorphic to the one-dimensional projective space:

  $$\mathcal{O}_d^d \cong P^1(\mathbb{Z}/d\mathbb{Z}).$$

  Namely, for $(a, b) \in U_d^d$ the quotient a/b only depends on the orbit

  $$\alpha_d^d(U(\mathbb{Z}/d\mathbb{Z}), (a, b)) \in \mathcal{O}_d^d$$

  through the state (a, b), and any two states with the same quotient lie in the same orbit.

- If d is a prime power ($d = p^\alpha$ with $p \in \mathbb{P}$), then as a set

  $$\mathcal{O}_d^d \cong P^1(\mathbb{Z}/d\mathbb{Z}) \times \mathbb{Z}/p^{\alpha-1}\mathbb{Z}.$$

- If d and e are relatively prime $\gcd(d,e)=1$), then for $f := de$

$$\mathcal{O}_f^f \cong \mathcal{O}_d^d \times \mathcal{O}_e^e.$$

The endomorphism $A_n : V_n \to V_n$,

$$(A_n f)(\omega) := |U(\mathbb{Z}/n\mathbb{Z})|^{-1} \sum_{x\in U(\mathbb{Z}/n\mathbb{Z})} f(x\cdot\omega)$$

is an orthogonal projection on the space $\bar{V}_n := A_n(V_n)$ of functions which are invariant on the α_n-orbits.

By Lemma 5 the subspaces $\bar{V}_n^d := A_n(V_n^d)$ have the form $\bar{V}_n^d = V_n^d \cap \bar{V}_n$. Their dimensions are $\dim(\bar{V}_n^d) = |\mathcal{O}_n^d|$, see (30).

By its definition (20), the vector $\mathbb{1}_n^d$ lies in $\bar{V}_n^d$, that is, $A_n(\mathbb{1}_n^d) = \mathbb{1}_n^d$.

We have $(A_n\delta_x, A_n\delta_y) = 0$ if $x, y \in U_n^d$ belong to different α_n-orbits and $(A_n\delta_x, A_n\delta_y) = |U(\mathbb{Z}/n\mathbb{Z})|^{-1}$ otherwise. So we find an orthonormal basis $\{\bar{e}_\omega \mid \omega\in\mathcal{O}_n^d\}$ of V_n^d by setting $\bar{e}_\omega := \sqrt{|U(\mathbb{Z}/n\mathbb{Z})|}\cdot A_n\delta_x$ for an arbitrary point $x\in U_n^d$ of the orbit $\omega\in\mathcal{O}_n^d$.

By the distributive law, the permutations τ_n^L and τ_n^R of Ω_n defined in (15) map α_n-orbits to α_n-orbits and thus induce permutations $\bar{\tau}_n^L$ and $\bar{\tau}_n^R$ of $\mathcal{O}_n$, leaving the subsets $\mathcal{O}_n^d$ invariant.

Similar to the last section, we denote the permutation representation of a permutation $\tau\in\mathcal{S}(\mathcal{O}_n)$ by $\bar{P}_\tau : \bar{V}_n \to \bar{V}_n$, and set $L_n := \bar{P}_{\bar{\tau}_n^L}$, $R_n := \bar{P}_{\bar{\tau}_n^R}$ for simplicity.

The doubly stochastic matrix T_n commutes with A_n, since

$$\begin{aligned} A_n P_{\tau_n^L} f((a,b)) &= |U(\mathbb{Z}/n\mathbb{Z})|^{-1} \sum_{x\in U(\mathbb{Z}/n\mathbb{Z})} f((x(a-b),xb)) \\ &= |U(\mathbb{Z}/n\mathbb{Z})|^{-1} \sum_{x\in U(\mathbb{Z}/n\mathbb{Z})} f(((xa)-(xb),xb)) = P_{\tau_n^L} A_n f((a,b)) \end{aligned}$$

and similarly for τ_n^R. Thus we may define $\bar{T}_n : \bar{V}_n \to \bar{V}_n$ by restriction $\bar{T}_n := T_n|_{\bar{V}_n}$. The relation

$$\bar{T}_n = \tfrac{1}{2}\left(L_n + R_n\right). \tag{31}$$

holds true.

The restrictions $\bar{T}_n^d := \bar{T}_n|_{\bar{V}_n^d}$ to the subspaces $\bar{V}_n^d$ define endomorphisms of these subspaces, so that

$$\bar{T}_n \cong \bigoplus_{d|n} \bar{T}_n^d \quad \text{on } \bar{V}_n \cong \bigoplus_{d|n} \bar{V}_n^d. \tag{32}$$

The matrix representations of $\bar{T}_n$ and $\bar{T}_n^d$ w.r.t. the basis vectors $\bar{e}_\omega$ are denoted by the same symbols.

By definition, the matrices $\bar{T}_n$ are doubly stochastic.

$\bar{T}_n^d$ arises from T_n^d by lumping together states in U_n^d belonging to the same orbit. Thus $\bar{T}_n^d$ is irreducible-aperiodic since T_n^d has that property (see Lemma 4).

Example. For the primes $n = d = 2$ resp. 3 one has

$$\bar{T}_2^2 = \frac{1}{2}\begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad \bar{T}_3^3 = \frac{1}{2}\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}. \tag{33}$$

in the basis corresponding to the enumeration $(1, \ldots, d, \infty)$ of the projective space $P^1(\mathbb{Z}/d\mathbb{Z})$.

By Perron-Frobenius, the eigenvalue one of $\bar{T}_n^d$ has algebraic multiplicity one, with eigenvector $\mathbb{1}_n^d$ and the spectral radius

$$\mathrm{sr}(\bar{T}_n^d - \bar{\Pi}_n^d) < 1, \tag{34}$$

$\bar{\Pi}_n^d := \Pi_n^d|_{\bar{V}_n^d}$ being the orthogonal projector on $\mathrm{span}(\mathbb{1}_n^d)$.

It was remarked in the beginning of this section that $\tilde{\chi}_m : \Omega_n \to \{0,1\}$ is constant on the α_n-orbits so that $A_n\tilde{\chi}_m = \tilde{\chi}_m$.

This implies that

$$\langle \chi_m \rangle_k = (\tilde{\chi}_m, (\bar{T}_n^n)^k \bar{v}_{n,0}) \tag{35}$$

with $\bar{v}_{n,k} := A_n v_{n,k}$, since by (24)

$$\begin{aligned} \langle \chi_m \rangle_k &= (\tilde{\chi}_m, (T_n^n)^k v_{n,0}) \\ &= (A_n\tilde{\chi}_m, (T_n^n)^k v_{n,0}) = (\tilde{\chi}_m, A_n(T_n^n)^k v_{n,0}) \end{aligned}$$

and $A_n T_n^n = T_n^n A_n$.

5 The Spectrum of $\bar{T}_n$

Now we will analyze $\bar{T}_n^d$ with more precision, in order to show that (unlike in the case of T_n^d) the spectral radius $\mathrm{sr}(\bar{T}_n^d - \bar{\Pi}_n^d) \le 1/\sqrt{2}$.

By the isomorphism $\mathcal{O}_n^d \to \mathcal{O}_d^d$ mentioned above it suffices to consider the special case $d = n$.

The first remark which will be of importance in the determination of the spectrum $\sigma(\bar{T}_n^d)$ is that the group of permutations of the set $\mathcal{O}_n^d$ generated by $\bar{\tau}^L$ and $\bar{\tau}^R$ is much smaller than the full permutation group $\mathcal{S}(\mathcal{O}_n^d)$.

Writing the representatives $(a,b) \in U_n^n$ of the orbits in the form of column vectors $\binom{a}{b}$, left addition $\tau_n^L : \binom{a}{b} \mapsto \binom{a+b}{b}$ is represented by the matrix $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$, and right addition $\tau_n^R : \binom{a}{b} \mapsto \binom{a}{a+b}$ is represented by the matrix $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$. So the subgroup of permutations generated by $\bar{\tau}_n^L$ and $\bar{\tau}_n^R$ is isomorphic to the group of matrices in $\mathrm{Mat}(2,\mathbb{Z}/n\mathbb{Z})$ generated by $\left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}1&0\\1&1\end{smallmatrix}\right)$.

Example. If n is a prime number, then we recover the group of Möbius transformations

$$z \mapsto \frac{az+b}{cz+d}, \qquad z \in \mathbb{P}^1(\mathbb{Z}/n\mathbb{Z}),$$

with $a,b,c,d \in \mathbb{Z}/n\mathbb{Z}$ and $ad-bc=1$.

Now we represent right addition τ^R in the form $\tau^R = \tau^I\tau^L\tau^I$, with the permutation $\tau^I(a,b) := (b,a)$, $(a,b) \in U_n^d$. For simplicity we introduce $\tau^M(a,b) := (a,-b)$ as a further permutation on U_n^d. As both of these order two permutations act on orbits, they induce permutations $\bar{\tau}^I$ and $\bar{\tau}^M$ on $\mathcal{O}_n^d$, and $\bar{\tau}^I\bar{\tau}^M = \bar{\tau}^M\bar{\tau}^I$ although $\tau^I\tau^M \neq \tau^M\tau^I$ on U_n^d for $d>2$:

$$\tau^I\tau^M(a,b) = (-b,a) \text{ and } \tau^M\tau^I(a,b) = (b,-a),$$

but $(b,-a) = -1\cdot(-b,a)$ belongs to the orbit through $(-b,a)$.

Moreover, we use the shorthands

$$L := \bar{P}_{\bar{\tau}^L}, \quad R := \bar{P}_{\bar{\tau}^R}, \tag{36}$$

$I := \bar{P}_{\bar{\tau}^I}$ and $M := \bar{P}_{\bar{\tau}^M}$ for the matrices of the permutation representations, and $\bar{T} \equiv \bar{T}_d^d$.

The following identities (depicted in Figure 3) turn out to be useful:

Lemma 7

$$J := MI = LR^{-1}L = L^{-1}RL^{-1} = R^{-1}LR^{-1} = RL^{-1}R = IM.$$

Proof. $MI = LR^{-1}L$ since

$$\tau^L\tau^{R^{-1}}\tau^L(a,b) = \tau^L\tau^{R^{-1}}(a+b,b) = \tau^L(a+b,-a) = (b,-a) = \tau^M\tau^I(a,b).$$

Similarly $IM = RL^{-1}R$.

Then we note that $IM = MI = (MI)^{-1}$ since $\bar{\tau}^I\bar{\tau}^M = \bar{\tau}^M\bar{\tau}^I$, and since both $\bar{\tau}^I$ and $\bar{\tau}^M$ are of order two. So $LR^{-1}L = L^{-1}RL^{-1}$ and $R^{-1}LR^{-1} = RL^{-1}R$. □

Lemma 8

$$\bar{T}^{-1} = 2\bar{T}^* - J.$$

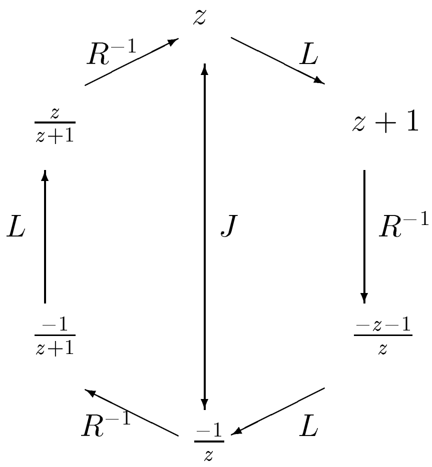

Figure 3: Relations between the Möbius transformations and the matrices of the permutation representation

Proof. We show that $2\bar{T}\bar{T}^* - \bar{T}J = \mathbb{1}$. Using $\bar{T} = \frac{1}{2}(L+R)$, and the orthogonality of the permutation matrices, we have

$$2\bar{T}\bar{T}^t = \tfrac{1}{2}(L+R)(L^{-1}+R^{-1}) = \mathbb{1} + \tfrac{1}{2}(LR^{-1}+RL^{-1}).$$

So we must prove that $\bar{T}J = \frac{1}{2}(LR^{-1}+RL^{-1})$. But

$$\begin{aligned} 2\bar{T}J &= LJ + RJ = L(L^{-1}RL^{-1}) + R(R^{-1}LR^{-1}) \\ &= RL^{-1} + LR^{-1}, \end{aligned}$$

using Lemma 7. □

Proposition 9 *If $t \in \mathbb{C}$ is an eigenvalue of $\bar{T}_n^d - \bar{\Pi}_n^d$, then $t \in (-1, -\frac{1}{2}] \cup [\frac{1}{2}, 1)$ or $|t| = 1/\sqrt{2}$.*

Proof. Let $f \in \bar{V}_n^d$, $(f,f) = 1$ be an eigenvector with eigenvalue t. Then from Lemma 8 we conclude that

$$2\bar{t} - t^{-1} = (f, J\,f). \tag{37}$$

Moreover, $-1 \leq (f, J\,f) \leq 1$ since J is a self-adjoint involution. Writing the l.h.s. of (37) in the form $\bar{t}(2 - |t|^{-2})$, one sees that $|t| = 1/\sqrt{2}$ if $t \notin \mathbb{R}$.

If $t \in \mathbb{R}$, one obtains from (37) $t \in [-1, -\frac{1}{2}] \cup [\frac{1}{2}, 1]$. On the other hand, we already remarked in (34) that by the Perron-Frobenius theorem the spectral radius $\mathrm{sr}(\bar{T}_n^d - \bar{\Pi}_n^d) < 1$. □

In order to proceed in the spectral analysis of $\bar{T}$, we introduce the operators

$$Y^+ := \tfrac{1}{3}(\mathbb{1} + R^{-1}L + L^{-1}R) = IY^+I \tag{38}$$

and

$$Y^- := \tfrac{1}{3}(\mathbb{1} + LR^{-1} + RL^{-1}) = JY^+J \tag{39}$$

which by Lemma 7 are ortho-projections. Geometrically Y^+ (Y^-) corresponds to the mean over the orbits generated by the order three transformation $(\bar{\tau}^R)^{-1}\bar{\tau}^L$ (resp. $\bar{\tau}^L(\bar{\tau}^R)^{-1}$).

Considering a pair of projectors Y^+, Y^- it is generally useful to introduce the operators

$$A := Y^+ - Y^- \qquad \text{and} \qquad B := \mathbb{1} - Y^+ - Y^-,$$

see Avron, Seiler and Simon [2]. These meet the relations

$$A^2 + B^2 = \mathbb{1}, \qquad AB + BA = 0$$

and $[A^2, Y^\pm] = [B^2, Y^\pm] = 0$.

In our case the non-normal operator $\bar{T}$ is related to the self-adjoint operator B by the following formula

Lemma 10

$$\left(2\bar{T} + \bar{T}^{-1}\right)^2 = 9B^2.$$

Proof. With the help of Lemma 7 and Lemma 8 one shows the identities

$$\bar{T} = \tfrac{3}{2}JY^+ - \tfrac{1}{2}J, \qquad \bar{T}^{-1} = 3Y^+J - 2J.$$

So

$$2\bar{T} + \bar{T}^{-1} = 3(JY^+ + Y^+J - J) = -3JB = -3BJ.$$

This implies the formula, since $J^2 = \mathbb{1}$. $\square$

In particular Lemma 10 shows that the algebraic and geometric multiplicity of an eigenvalue t of the Markov transition matrix $\bar{T}$ coincide if $t \neq \pm 1/\sqrt{2}$ (since $2\bar{T} + \bar{T}^{-1}$ is self-adjoint).

We now split the Hilbert space $\bar{V} \equiv \bar{V}_n^d$ on which $\bar{T}$ acts into the orthogonal direct sum

$$\bar{V} = \bar{V}^{\mathrm{ker}} \oplus \bar{V}^{\mathrm{ran}} \tag{40}$$

with

$$\bar{V}^{\mathrm{ker}} := \ker(Y^+) \cap \ker(Y^-), \qquad \bar{V}^{\mathrm{ran}} := \mathrm{ran}(Y^+) + \mathrm{ran}(Y^-).$$

This splitting induces a decomposition of our operators B and $\bar{T}$.

We begin with the simpler piece. By definition of B, the restriction of B to $\bar{V}^{\ker}$ is the identity operator. By Lemma 10 an eigenvalue ± 1 of B corresponds to an eigenvalue $\pm\frac{1}{2}$ or ± 1 of $\bar{T}$.

We already know that the multiplicity of the eigenvalue one of $\bar{T}$ is one and that -1 does not occur in its spectrum ($\bar{T}$ being an irreducible-aperiodic doubly stochastic operator). So $\bar{V}^{\ker}$ equals the orthogonal sum of the eigenvalue $\frac{1}{2}$ and $-\frac{1}{2}$ subspaces:

$$\bar{V}^{\ker} = \ker(\bar{T} - \tfrac{1}{2}\mathbb{1}) \oplus \ker(\bar{T} + \tfrac{1}{2}\mathbb{1}) \qquad (41)$$

Both of them have approximately dimension $d/6$:

Proposition 11 *Let $d \in \mathbb{N}$ be a prime number. Then the dimensions*

$$D^{\pm\frac{1}{2}} := \dim\left(\ker(\bar{T} \mp \tfrac{1}{2}\mathbb{1})\right)$$

of the eigenvalue $\pm\frac{1}{2}$ subspaces of $\bar{T} \equiv \bar{T}_d^d$ equal

- *for $d = 2$, $D^{+\frac{1}{2}} = D^{-\frac{1}{2}} = 1$,*
- *for $d = 3$, $D^{+\frac{1}{2}} = 1$ whereas $D^{-\frac{1}{2}} = 0$,*
-

d (mod 12)	$D^{+\frac{1}{2}}$	$D^{-\frac{1}{2}}$
1	$\frac{d-1}{6} - 1$	$\frac{d-1}{6}$
5	$\frac{d+1}{6}$	$\frac{d+1}{6} + 1$
7	$\frac{d-1}{6}$	$\frac{d-1}{6} - 1$
11	$\frac{d+1}{6} + 1$	$\frac{d+1}{6}$

Proof. The statements for the primes 2 and 3 follow from direct inspection of the matrices (33).

So we may assume $2 \nmid d$ and $3 \nmid d$. Lemma 8 implies that the involution J, restricted to the $\bar{T}$-eigenspaces for eigenvalues $\pm\frac{1}{2}$, equals ∓ 1. So

$$\ker\left(\bar{T} \mp \tfrac{1}{2}\mathbb{1}\right) = \bar{V}^{\ker} \cap \ker\left(J \pm \mathbb{1}\right).$$

By the relation (39) between Y^- and Y^+ this eigenspace is also characterized by

$$\ker\left(\bar{T} \mp \tfrac{1}{2}\mathbb{1}\right) = \ker(Y^+) \cap \ker\left(J \pm \mathbb{1}\right).$$

So we count the dimension of the r.h.s., beginning with $\ker\left(J \pm \mathbb{1}\right)$.

1) The involution J corresponds to the permutation $z \mapsto -1/z$ on $P^1(\mathbb{Z}/d\mathbb{Z})$. This permutation has

- for $d = -1 \pmod 4$ no fixed points
- for $d = 1 \pmod 4$ the two fixed points $\pm\sqrt{-1} \in P^1(\mathbb{Z}/d\mathbb{Z})$,

since the Jacobi symbol equals $\left(\frac{-1}{d}\right) = (-1)^{\frac{1}{2}(d-1)}$.

The eigenvalue equation $J\varphi = \mu\varphi$ gives one independent equation per pair $(z, -1/z)$ with $z \neq -1/z$. If and only if $\mu = -1$, the fixed points $z = -1/z$ give additional equations (namely $\varphi(z) = 0$). So we obtain the following numbers of independent equations:

$d \pmod{12}$	$\mu = -1$	$\mu = 1$
1	$\frac{d+1}{2} + 1$	$\frac{d+1}{2} - 1$
5	$\frac{d+1}{2} + 1$	$\frac{d+1}{2} - 1$
7	$\frac{d+1}{2}$	$\frac{d+1}{2}$
11	$\frac{d+1}{2}$	$\frac{d+1}{2}$

2) Next we determine $\dim(\ker(Y^+))$. This equals the dimension $d+1$ of our Hilbert space $\bar{V}$, minus the number of orbits of the permutation $z \mapsto -1-1/z$ on $P^1(\mathbb{Z}/d\mathbb{Z})$. Since by quadratic reciprocity the Jacobi symbol

$$\left(\frac{-3}{d}\right) = (-1)^{\frac{1}{2}(d-1)}\left(\frac{3}{d}\right) = \left(\frac{d}{3}\right),$$

this permutation has

- for $d = -1 \pmod 6$ no fixed points. Thus every orbit consists of three points, and we have $\frac{d+1}{3}$ orbits.
- for $d = 1 \pmod 6$ the two different fixed points $-\frac{1}{2}(1 + \sqrt{-3})$ and $-\frac{1}{2}(1 - \sqrt{-3})$ corresponding to the roots of $z^2 + z + 1$ in $P^1(\mathbb{Z}/d\mathbb{Z})$. So in that case we have $\frac{d-1}{3} + 2$ orbits.

However, the set of equations obtained from the condition $\varphi \in \ker(Y^+)$ is not independent from set of equations obtained by the eigenvalue equation $J\varphi = \mu\varphi$ if $\mu = -1$. Namely we may obtain from $J\varphi = -\varphi$ that φ has zero mean value ($\sum_{z \in P^1(\mathbb{Z}/d\mathbb{Z})} \varphi(z) = 0$), a property which already follows from $Y^+\varphi = 0$.

So $\varphi \in \ker(Y^+)$ gives us at most

$d \pmod{12}$	$\mu = -1$	$\mu = 1$
1	$\frac{d-1}{3} + 1$	$\frac{d-1}{3} + 2$
5	$\frac{d+1}{3} - 1$	$\frac{d+1}{3}$
7	$\frac{d-1}{3} + 1$	$\frac{d-1}{3} + 2$
11	$\frac{d+1}{3} - 1$	$\frac{d+1}{3}$

independent additional equations.

3) Subtracting the above two lists from the dimension $d+1$ of the Hilbert space $\bar{V}$, we see that the dimensions $D^{\pm\frac{1}{2}}$ are at least as large as stated in our lemma.

To obtain the upper bounds for $D^{\pm\frac{1}{2}}$, we compare dimensions in (41). As calculated above,

$$\dim(\mathrm{ran}(Y^+)) = (d+1)/3 \quad \text{if} \quad d = -1 \,(\mathrm{mod}\, 6) \tag{42}$$

and

$$\dim(\mathrm{ran}(Y^+)) = (d-1)/3 + 2 \quad \text{if} \quad d = 1 \,(\mathrm{mod}\, 6). \tag{43}$$

The same holds true for the projector $Y^- = JY^+J$.

However, there is exactly one relation between the equations characterizing $\ker(Y^+)$ and $\ker(Y^-)$.

Each $z \in P^1(\mathbb{Z}/d\mathbb{Z})$ appears exactly once in both sets of equations, which are of the form

$$\varphi(z) + \varphi(-1-1/z) + \varphi(-1/(z+1)) = 0 \tag{44}$$

for $\varphi \in \ker(Y^+)$ respectively

$$\varphi(z) + \varphi(1-1/z) + \varphi(-1/(z-1)) = 0 \tag{45}$$

for $\varphi \in \ker(Y^-)$. All coefficients of these equations equal $+1$. So the sum of the equations (44) minus the sum of the equations (45) equals zero. But this is the only relation between eqs. (44) and (45).

So by (42) and (43) the dimension of

$$\bar{V}^{\mathrm{ker}} = \ker(Y^+) \cap \ker(Y^-)$$

equals

$$\dim(\bar{V}) - \dim(\mathrm{ran}(Y^+)) - \dim(\mathrm{ran}(Y^-)) + 1 = d + 2 - 2\dim(\mathrm{ran}(Y^+))$$

and thus $D^{+\frac{1}{2}} + D^{-\frac{1}{2}}$, proving the lemma. □

For general integers d the dimensions $D^{\pm\frac{1}{2}}$ may be determined along the same lines.

6 Ramanujan Graphs

Now we turn to the part of the spectrum of $\bar{T}$ which belongs to the subspace $\bar{V}^{\mathrm{ran}}$ in the orthogonal decomposition (40) of the Hilbert space $\bar{V} \equiv \bar{V}_d^d$. The constant eigenfunction with eigenvalue 1 belongs to that subspace.

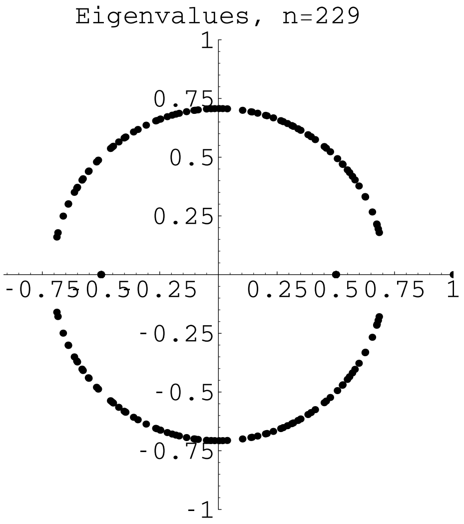

Figure 4: The spectrum of the reduced Markov transition matrix $\bar{T}$ for the 50th prime $n = d = 229$.

The dimension of the co-dimension one subspace $\bar{V}^{\mathrm{ran},\perp} \subset \bar{V}^{\mathrm{ran}}$ of functions with zero mean is even and follows from (40) and (41) which give

$$\dim(\bar{V}^{\mathrm{ran},\perp}) = \dim(\bar{V}) - 1 - D^{+\frac{1}{2}} - D^{-\frac{1}{2}}.$$

So by Prop. 11 it is 0 for $d = 2$, and 2 for $d = 3$. For $d \in \mathbb{P}$ one has $\dim(\bar{V}^{\mathrm{ran},\perp}) = \frac{2}{3}(d-1)+2$ if $d = 1 \pmod 6$ and $\frac{2}{3}(d+1)-2$ if $d = -1 \pmod 6$.

We know from Prop. 9 that the non-real eigenvalues of $\bar{T}$, which are all associated to its restriction to $\bar{V}^{\mathrm{ran},\perp}$, have absolute value $1/\sqrt{2}$.

We would like to show that there are no real eigenvalues t of $\bar{T}|_{\bar{V}^{\mathrm{ran},\perp}}$, since these would in general enlarge the spectral radius. For the first 50 primes we checked this property, see Figure 4.

By Lemma 10 there do not exist such additional real eigenvalues if

$$\mathrm{spec}\,(B|_{\bar{V}^{\mathrm{ran},\perp}}) \subset \left(-\frac{\sqrt{8}}{3}, \frac{\sqrt{8}}{3}\right).$$

Without restriction to the subspace this corresponds to the property

$$\mathrm{spec}\,(B) \subset \{-1\} \cup \left(-\frac{\sqrt{8}}{3}, \frac{\sqrt{8}}{3}\right) \cup \{1\} \tag{46}$$

of $B = \mathbb{1} - Y^+ - Y^-$. We will now relate relation (46) to the so-called Ramanujan property of certain graphs.

The key observation is that by (36) the definition (38), (39) of the ortho-projectors $Y^\pm$ is related to the action of the matrices $M_\pm \in \mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$

$$M_+ := \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad M_- := \begin{pmatrix} -1 & 1 \\ -1 & 0 \end{pmatrix}. \tag{47}$$

These group elements of order three are conjugated by

$$M_+^2 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} M_- \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

and $M_+ \neq M_-$ iff $d > 2$. M_+ corresponds to the transformation $R^{-1}L$, whereas M_- corresponds to RL^{-1}.

Of course, M_+ and M_- also act by left transformations on the group $\mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$. For $d > 1$ the orbits of these actions are of size three. For $d > 2$ the M_+-orbit and the M_--orbit through $g \in \mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$ have only g in common.

Definition 12 *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 $G = (V, E)$. 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 $g \in \mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$.

For $d > 2$ (and we will henceforth consider only that case) the graph G is *three-regular*, that is, any vertex has three adjacent edges, and it is *connected*, that is, any two vertices in V are connected by a chain of edges in E. E is then naturally isomorphic to $\mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$. Since V_- and V_+ are disjoint, G is bipartite.

Example. For $d = 3$ the order of the group $\mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$ is 24. So $|V| = 2 \times 8 = 16$. In this case G can be visualized as follows: Attach two vertices at each corner of a cube, one inside, and one outside the cube. Connect vertices along the edges of the cube, changing between inside and outside along three edges of maximal distance, see Figure 5. By using $\mathrm{PSL}(2, \mathbb{Z}/3\mathbb{Z})$ instead of $\mathrm{SL}(2, \mathbb{Z}/3\mathbb{Z})$, we obtain the graph of the cube.

Now we consider the Laplacian Δ of the graph G.

We remind the reader of the definition of the *Laplacian Δ of a graph* $G = (V, E)$.

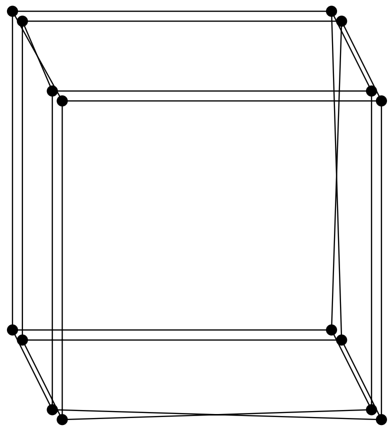

Figure 5: The graph G for the group $\mathrm{SL}(2, \mathbb{Z}/3\mathbb{Z})$.

Consider the Hilbert space $\mathcal{H} := \mathcal{H}_0 \oplus \mathcal{H}_1$ with $\mathcal{H}_0 := l^2(V)$ (with counting measure) and $\mathcal{H}_1$ being isomorphic to $l^2(E)$. More precisely we double the unoriented edges by setting $\mathbf{E} := \{(v,w) \mid \{v,w\} \in E\}$ and consider the subspace

$$\mathcal{H}_1 := \{f \in l^2(\mathbf{E}) \mid f((w,v)) = -f((v,w))\}$$

with inner product $(f,g) := \frac{1}{2}\sum_{\mathbf{e}\in\mathbf{E}} \bar{f}(\mathbf{e})g(\mathbf{e})$.

Then the adjoint of

$$d : \mathcal{H}_0 \to \mathcal{H}_1, \qquad df((v,w)) := f(w) - f(v)$$

equals

$$d^* : \mathcal{H}_1 \to \mathcal{H}_0, \qquad d^*g(v) = -\sum_{(v,w)\in\mathbf{E}} g((v,w)).$$

As usual one defines $\Delta : \mathcal{H} \to \mathcal{H}$ by

$$\Delta := d^*d + dd^*,$$

so that $\Delta = \Delta_0 \oplus \Delta_1$ with

$$\Delta_0 f(v) = \sum_{\{v,w\}\in E} (f(v) - f(w))$$

and

$$\Delta_1 g((v,w)) = \sum_{(v,x)\in\mathbf{E}} g((v,x)) + \sum_{(x,w)\in\mathbf{E}} g((x,w)).$$

Thus we are in a supersymmetric situation (see, e.g., [7]) and, apart from zero eigenvalues, the spectra of Δ_0 and Δ_1 coincide, including multiplicities.

In the case of a k-regular graph G (that is, a graph whose vertices have degree k), $\Delta_0 = k\mathbb{1} - A$, A being the adjacency matrix of G. If G is finite, then k is an eigenvalue (corresponding to the constant eigenfunction) of A, and we denote by $\mu \geq 0$ the absolute value of the next to largest eigenvalue (in absolute value) of A.

If G is bipartite, then the eigenvalues of A are symmetrically distributed around 0, so that also $-k$ is an eigenvalue in the finite case. Thus we denote by $\mu_b \geq 0$ the absolute value of the third to largest eigenvalue (in absolute value) of A.

Definition 13 *A k-regular graph G is called* **Ramanujan** *if*

$$\mu \leq 2\sqrt{k-1}.$$

A bipartite k-regular graph G is called **bipartite Ramanujan** *if*

$$\mu_b \leq 2\sqrt{k-1}. \tag{48}$$

Conjecture. The graphs G intoduced in Def. 12 for $\mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})$ are bipartite Ramanujan.

Proposition 14 *If the above conjecture holds true, then instead of (34) we obtain the optimal estimate*

$$\mathrm{sr}\left(\bar{T}_n^d - \bar{\Pi}_n^d\right) \leq 1/\sqrt{2},$$

for the spectral radius of the reduced transition matrix. If in addition the inequality (48) is strict, then for some $C > 0$

$$\left\| \left(\bar{T}_n^d - \bar{\Pi}_n^d\right)^k \right\| \leq C \cdot 2^{-k/2} \qquad (k \in \mathbb{N}). \tag{49}$$

Proof. The spectrum of B is a subset of the spectrum of its lift $\bar{B}$ to the group ring $\mathbb{C}[\mathrm{SL}(2, \mathbb{Z}/d\mathbb{Z})]$. Now $3\bar{B} = 3\mathbb{1} - \Delta_1$, with Δ_1 being the edge Laplacian for the graph G of Def. 12. So (apart from multiplicity of the eigenvalue 3) the spectrum of $3\bar{B}$ equals the spectrum of the adjacency matrix $A = 3\mathbb{1} - \Delta_0$. This implies that (46) holds true if the three-regular graph G is Ramanujan.

We remarked that as a consequence of Lemma 10 $\bar{T}$ is semi-simple if no eigenvalue equals $\pm 1/\sqrt{2}$. This is the case under our assumption on the spectrum of the graph. But a semi-simple matrix is normal in some metric, and the norm of a power of a normal matrix equals the power of the norm. $\square$

Ramanujan graphs are in a sense optimal, since the estimate $\mu \leq c$ is always violated for large k-regular graphs if $c < 2\sqrt{k-1}$. They have number-theoretic and engineering applications.

In the articles [3] of Chiu, [17] by Lubotzky, Phillips and Sarnak, [18] by Margulis, [22, 23] by Pizer and the books [24] by Sarnak and [16] by Lubotzky one finds different constructions leading to Ramanujan graphs. Ref. [26] by Venkov and Nikitin is a general survey.

However, our family of graphs described in Def. 12 does not fall in one of these known families of Ramanujan graphs.

7 Estimating Expectations

Our aim is to estimate expectations $\langle f\rangle_k$, beginning with $\langle \chi_m\rangle_k$.

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

$$\chi_{m_1} \cdot \chi_{m_2} = \chi_m \qquad \text{with } m := \frac{m_1 m_2}{\gcd(m_1, m_2)} \tag{50}$$

being the least common multiple of m_1 and m_2. This implies a rule for multiplying the $c_m = 1 - 2\chi_m$. In particular we are interested in products of the functions $\lambda_{p,r} := \prod_{l=1}^{r} c_{p^l} : \mathbb{N} \to \{-1, 1\}$, $p \in \mathbb{P}$, $r \in \mathbb{N}_0$ which approximate the function λ_p defined in (8) in the sense that $\lambda_{p,r}(m) = \lambda_p(m)$ for $m < p^{r+1}$.

Lemma 15

$$\lambda_{p,r} = 1 + 2\sum_{l=1}^{r}(-1)^l \chi_{p^l}.$$

Proof. $\lambda_{p,1} = c_p = 1 - 2\chi_p$ and

$$\begin{aligned}\lambda_{p,r+1} &= \lambda_{p,r} \cdot c_{p^{r+1}} = \left(1 + 2\sum_{l=1}^{r}(-1)^l \chi_{p^l}\right)(1 - 2\chi_{p^{r+1}}) \\ &= 1 + 2\sum_{l=1}^{r}(-1)^l \chi_{p^l} - 2\chi_{p^{r+1}}\left(1 + 2\sum_{l=1}^{r}(-1)^l\right),\end{aligned}$$

using (50). But $1 + 2\sum_{l=1}^{r}(-1)^l = (-1)^r$. □

The existence of thermodynamic limits follows from the Perron-Frobenius Theorem:

Proposition 16 *For $m \in \mathbb{N}$*

$$\langle \chi_m\rangle_\infty = \left(m \prod_{p\in\mathbb{P}, p|m} (1 + 1/p)\right)^{-1}, \tag{51}$$

for $p \in \mathbb{P}$, $r \in \mathbb{N}$

$$\langle \lambda_{p,r} \rangle_\infty = \frac{p^2+1-2(-1/p)^{r-1}}{(p+1)^2} \quad \text{and} \quad \langle \lambda_p \rangle_\infty = \frac{p^2+1}{(p+1)^2}. \tag{52}$$

Furthermore, for any set $\{p_1, \ldots, p_s\} \subset \mathbb{P}$ *of primes and any numbers* $r_1, \ldots, r_s \in \mathbb{N} \cup \infty$

$$\left\langle \prod_{i=1}^{s} \lambda_{p_i,r_i} \right\rangle_\infty = \prod_{i=1}^{s} \langle \lambda_{p_i,r_i} \rangle_\infty . \tag{53}$$

Finally $\langle \lambda \rangle_\infty = 0$.

Proof. The limit (51) (defined in (9)) exists, since by (35)

$$\langle \chi_m \rangle_k = (\tilde{\chi}_m, (\bar{T}_m^m)^k \bar{v}_{m,0}),$$

and the doubly stochastic matrix $\bar{T}_m^m$ is irreducible-aperiodic. This matrix acts on the $|\mathcal{O}_m^m| = m \prod_{p \in \mathbb{P}, p|m} (1+1/p)$–dimensional vector space $\bar{V}_m^m$ (see (30)), and $\tilde{\chi}_m$ is a one-dimensional projection on $\bar{V}_m^m$. This implies (51).

Then the first formula of (52) follows from Lemma 15. This also implies that

$$\lambda_{p,2r-1} \le \lambda_p \le \lambda_{p,2r},$$

whence the second part of (52). The independence (53) of the expectations in the thermodynamic limit is a consequence of the fact that the dimension $|\mathcal{O}_m^m|$ of $\bar{V}_m^m$ is a multiplicative function of m.

$\langle \lambda \rangle_\infty = 0$, since by (52) and (53) for $l \in \mathbb{N}$

$$\ln \left\langle \prod_{p \in \mathbb{P}, p \le l} \lambda_p \right\rangle_\infty = \sum_{p \in \mathbb{P}, p \le l} \ln \left(\frac{p^2+1}{(p+1)^2} \right)$$

which diverges to $-\infty$ as $l \to \infty$, since

$$\frac{p^2+1}{(p+1)^2} \sim 1 - \frac{1}{2p},$$

and $\sum_{p \in \mathbb{P}} p^{-1} = \infty$. □

The estimate

$$|\langle \chi_m \rangle_k - \langle \chi_m \rangle_\infty| \le C(m) 2^{-k/2}$$

for the deviation (27) from the thermodynamic limit would follow from the validity of (49).

A The Lewis Equation

In [5] we showed that the *free energy*

$$F(\beta) := \lim_{k\to\infty} \frac{-1}{\beta k} \ln(Z_k(\beta))$$

of the Number-Theoretical Spin Chain equals for $0 < \beta < 2$

$$F(\beta) = -\ln(\lambda(\beta))/\beta,$$

where $\lambda(\beta)$ is the largest eigenvalue of a transfer operator $\tilde{C}(\beta)$ (by (4) $F(\beta) = 0$ for $\beta \geq 2$). This operator on $l^2(\mathbb{N}_0)$ has matrix elements

$$\tilde{C}(\beta)_{m,r} = (-1)^r 2^{-\beta-m-r} \left[\binom{-\beta-m}{r} + \sum_{s=0}^{m} 2^s \binom{m}{s} \binom{-\beta-m}{r-s} \right], \tag{54}$$

($m, r \in \mathbb{N}_0$), with the binomial coefficients $\binom{a}{b} = (\prod_{i=0}^{b-1}(a-i))/b!$, $a \in \mathbb{R}$, $b \in \mathbb{N}_0$, and $\binom{a}{b} = 0$ if $b < 0$.

Proposition 17 *The eigenvalue $\lambda(\beta)$ coincides with the largest eigenvalue λ of the Lewis three-term functional equation*

$$\lambda \cdot \psi(x) = \psi(x+1) + x^{-\beta}\psi(1+1/x). \tag{55}$$

Proof. For analytical questions (in particular existence of the Perron-Frobenius eigenvalue of multiplicity one) we refer to [14] resp. [5]. First we transform (55) by substituting $1/x$ for x and then dividing through x^β:

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

The r.h.s. coincides with the r.h.s. of (55). Thus

$$\psi(w) = w^{-\beta}\psi(1/w), \tag{56}$$

and we use (56) to transform the r.h.s. of (55):

$$K_\beta \psi = \lambda \cdot \psi \tag{57}$$

with $K_\beta : L^2((0,1)) \to L^2((0,1))$,

$$K_\beta \psi(x) := (x+1)^{-\beta} \left(\psi\left(\frac{x}{x+1}\right) + \psi\left(\frac{1}{x+1}\right) \right).$$

Expanding ψ around 1 in the form $\psi(w) = \sum_m a_m(1-w)^m$, we obtain

$$(x+1)^{-\beta}\psi\left(\frac{x}{x+1}\right) = \sum_{m=0}^{\infty} a_m(x+1)^{-\beta-m} \tag{58}$$
$$= \sum_{m=0}^{\infty} a_m 2^{-\beta-m}\sum_{r=0}^{\infty}(-1)^r\binom{-\beta-m}{r}2^{-r}(1-x)^r,$$

since

$$(x+1)^\alpha = 2^\alpha(1+\tfrac{1}{2}(x-1))^\alpha = 2^\alpha\textstyle\sum_{r=0}^{\infty}\binom{\alpha}{r}2^{-r}(x-1)^r.$$

Similarly

$$(x+1)^{-\beta}\psi\left(\frac{1}{x+1}\right) = \sum_{m=0}^{\infty} a_m x^m(x+1)^{-\beta-m} \tag{59}$$
$$= \sum_{m=0}^{\infty} a_m 2^{-\beta-m}\sum_{t=0}^{\infty}(-1)^{t+s}\sum_{s=0}^{m}\binom{m}{s}\binom{-\beta-m}{t}2^{-t}(1-x)^{t+s},$$

since

$$x^m = \sum_{s=0}^{m}\binom{m}{s}(x-1)^s.$$

The sum of (58) and (59) corresponds to (54). This proves the claim, since by the positivity of the a_m (see [5]) ψ is positive, too. $\square$

For $\beta \in \mathbb{Z} - \mathbb{N}$ the operator $\tilde{C}(\beta)$ leaves the subspace $\{f \in l^2(\mathbb{N}_0) \mid f(m) = 0 \text{ for } m > -\beta\}$ invariant, so that we obtain polynomial solutions of degree $|\beta| + 1$ of (57). The corresponding (eigenvalue one) solutions of

$$\psi(x) = \psi(x+1) + (x+1)^{2(k-1)}\psi(x/(x+1))$$

are called *period polynomials* and are related to the cusp forms of weight $2k$ of the modular group, see Lewis and Zagier [15].

Mayer showed in [19, 20] that the Selberg zeta function

$$Z_{\mathrm{SL}(2,\mathbb{Z})}(s) = \prod_{\substack{\{\gamma\}\in\mathrm{SL}(2,\mathbb{Z})\\ \text{primitive}}}\prod_{m=0}^{\infty}\left(1-\det(\gamma)^m\mathcal{N}(\gamma)^{-s-m}\right) \qquad (\mathrm{Re}(s) > 1)$$

with $\mathcal{N}(\gamma) = \left(\frac{1}{2}T(\gamma) + \sqrt{(T(\gamma)/2)^2 - \det(\gamma)}\right)^2$ can be written in the form

$$Z_{\mathrm{SL}(2,\mathbb{Z})}(s) = \det(\mathbb{1} - \mathcal{L}_s)\cdot\det(\mathbb{1} + \mathcal{L}_s),$$

$\mathcal{L}_s : A_\infty(D) \to A_\infty(D)$ being the transfer operator of the Gauss map:

$$\mathcal{L}_s f(x) := \sum_{n=1}^{\infty} (n+z)^{-2s} f\left(\frac{1}{n+z}\right).$$

Here $A_\infty(D)$ denotes the Banach space of functions holomorphic on the disk $D := \{z \in \mathbb{C} \mid |z-1| < 3/2\}$ and continuous on $\bar{D}$.

Since a fixed point f of $\mathcal{L}_s$ gives rise to a solution $\psi(x) := f(x-1)$ of the Lewis equation, we notice a relation with the Selberg zeta function.

References

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

[2] Avron, J., Seiler, R., Simon, B.: The Index of a Pair of Projections. J. Funct. Anal. **120**, 220–237 (1994)

[3] Chiu, P.: Cubic Ramanujan Graphs. Combinatorica **12**, 275–285 (1992)

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

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

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

[7] Cycon, H.L., Froese, R.G., Kirsch, W., Simon, B.: Schrödinger Operators. Texts and Monographs in Physics. Berlin: Springer 1987

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

[9] Guerra, F., Knauf, A.: Free Energy and Correlations of the Number-Theoretical Spin Chain. SFB 288 Preprint Nr. 191 (`http://www-sfb288.math.tu-berlin.de`) (1996)

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

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

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

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

[14] Lewis, J.B.: Spaces of holomorphic functions equivalent to the even Maass cusp forms. Invent. Math. **127**, 271–306 (1997)

[15] Lewis, J.B., Zagier, D.: Period functions for Maass wave forms. MPI 96 - 112 Preprint (`http://www.mpim-bonn.mpg.de`) (1996)

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

[17] Lubotzky, A., Phillips, R., Sarnak, P.: Ramanujan Graphs. Combinatorica **8**, 261–277 (1988)

[18] Margulis, G.A.: Explicit Group-Theoretical Constructions of Combinatorial Schemes and their Application to the Design of Expanders and Concentrators. Problemy Peredachi Informatsii **24**, 51–60 (1988)

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

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

[21] Odlyzko, A.M., te Riele, H.J.J.: Disproof of Mertens conjecture. J. f. d. reine und angewandte Math. **357**, 138–160 (1985)

[22] Pizer, A.: An Algorithm for Computing Modular Forms on $\Gamma_0(N)$. Journal of Algebra **64**, 340–390 (1980)

[23] Pizer, A.: Ramanujan Graphs and Hecke Operators. Bull. AMS (New Series) **23**, 127–137 (1990)

[24] Sarnak, P.: Some applications of modular forms. Cambridge, New York: Cambridge University Press 1990

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

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