

Native ultrametricity of sparse random ensembles

V. Avetisov^{1,2}, P. L. Krapivsky³ and S. Nechaev^{4,5}

¹*N. N. Semenov Institute of Chemical Physics of the Russian Academy of Sciences, 119991, Moscow, Russia*

²*Department of Applied Mathematics, National Research University Higher School of Economics, 101000, Moscow, Russia*

³*Physics Department, Boston University, Boston MA 02215, USA*

⁴*Université Paris-Sud/CNRS, LPTMS, UMR8626, 91405 Orsay, France*

⁵*P. N. Lebedev Physical Institute of the Russian Academy of Sciences, 119991, Moscow, Russia*

(Dated: October 18, 2015)

We investigate the eigenvalue density in ensembles of large sparse Bernoulli random matrices. Analyzing in detail the spectral density of ensembles of linear subgraphs, we discuss its ultrametric nature and show that near the spectrum boundary, the tails of the spectral density exhibit a Lifshitz singularity typical for Anderson localization. We pay attention to an intriguing connection of the spectral density to the Dedekind η -function. We conjecture that ultrametricity emerges in rare-event statistics and is inherit to generic complex sparse systems.

I. INTRODUCTION

A. Some notions about ultrametricity

Generally, a metric space is a set of elements equipped by pairwise distances. Generically, the metric $d(x, y)$ meets three requirements: i) non-negativity, $d(x, y) > 0$ for $x \neq y$, and $d(x, y) = 0$ for $x = y$, ii) symmetry, $d(x, y) = d(y, x)$, and iii) the triangle inequality, $d(x, z) \leq d(x, y) + d(y, z)$. The concept of ultrametricity is related to a special class of metrics, obeying the strong triangular inequality, $d(x, z) \leq \max\{d(x, y), d(y, z)\}$ (this condition allows the existence of acute isosceles and equilateral triangles only).

The strong triangle inequality radically changes the properties of spaces in comparison with our intuitive expectations based of Euclidean geometry. In ultrametric spaces, the sum of distances does not exceed the largest summand, i.e. the Archimedean principle does not hold. Few counterintuitive properties are related to ultrametric balls: every point inside a ball is its center; the radius of a ball is equal to its diameter; there are no balls that overlap only partially, if two balls have a common point, then one of them is completely embedded into the other. The last property implies that any ultrametric ball can be divided into a set of smaller balls, each of them can be divided into even smaller ones, etc. In other words, an ultrametric space resembles not a line of segments, but a branching tree of hierarchically nested balls (“basins”). The terminal points of a tree, constituting the tree boundary, also obeys the strong triangle inequality: for any two points, A and B, the distance to the root of the minimal subtree to which these points belong, determines the distance between A and B. As a result, ultrametric geometry, in general, fixes taxonomic (i.e. hierarchical) tree-like relationships between the elements and, speaking figuratively, is closer to Lobachevsky geometry, rather to the Euclidean one.

Ultrametricity is tightly linked to the p -adic norm which originally arose in the context of number theory [1]. The role of ultrametric spaces and p -adic numbers in number theory [2, 3] and in other branches of mathematics, such as p -adic analysis [4] and non-Archimedean analytic geometry [5], is growing. More recently, there have been numerous attempts to apply p -adic numbers and ultrametric spaces to various branches of physics ranging from string theory to inflationary cosmology (see e.g. [6–11]).

Perhaps most naturally ultrametric spaces emerge in the realm of complex systems. Indeed, the ultrametric description is inherently multi-scale, thus it is not surprising that ultrametricity provides a fertile framework for describing complex systems characterized by a large number of order parameters. For instance, signatures of ultrametricity were discovered in the context of spin glasses (see [12] for an earlier review). Spin glasses are characterized by a huge number of metastable states and the equilibration in spin glasses proceeds via tree-like splitting of the phase space into hierarchically nested domains. When the temperature decreases, the most distant groups of spin-glass states get factorized first. Then, each of these groups splits into a number of subgroups, etc. The scales of hierarchically nested sets of equilibrated states satisfy the strong triangle inequality, i.e. the multi-scale splitting of spin-glass phases obey ultrametric relations.

The physical origin behind the emergence of ultrametricity in spin-glass systems is quite general. In a complex system with an extremely large number of metastable states, the phase trajectories cannot totally explore all states: they cover only a tiny part of phase space almost without returns to already visited regions. For this reason, the factorization of spin-glass phases is similar to random branching in a high-dimensional (infinitely-dimensional in the thermodynamic limit) space. The factorization of spin-glass phases is specified only by branching points because once two branches are dispersed in a high-dimensional space, they never join. The ultrametric ansatz [13, 14] proposed to describe the spin-glasses by infinite (in the thermodynamic limit) number of order parameters, was proven in [15] to be an exact ground state of the Sherrington-Kirkpatrick model [16].

Shortly after the formulation of the ultrametric ansatz in the context of spin glasses, similar ideas were proposed for the description of native states of protein molecules [17, 18]. Proteins are highly frustrated polymer systems characterized by rugged energy landscapes of very high dimensionality. The ultrametricity of proteins implies a multi-scale representation of the protein energy landscape by means of ranging the local minima into hierarchically embedded basins of minima. Ultrametricity follows from the conjecture that the equilibration time within the basin is significantly smaller than the time to escape the basin. As a result, the transitions between local minima obey the strong triangle inequality. In such an approximation, dynamics on a complex energy landscape can be modelled via a jump-like random process, propagating in an ultrametric space of states. The approximation of the protein energy landscape by a self-similarly branching tree of basins has been shown to be very fruitful for describing a large body of experimentally justified features of protein conformational dynamics and protein fluctuation mobility at temperatures covering very wide range from room temperatures up to the deeply frozen states [19].

With respect to ultrametricity and the self-similarity of protein energy landscapes, it is important to keep in mind that proteins are functional systems. Proteins precisely operate at the atomic level using sparse-event statistics, which is a bridge between non-local protein dynamics in high-dimensional conformational space and low-dimensional movements of particular atomic units related to protein functionality.

For completeness, let us mention, that recently the ultrametric concepts based on generalized Parisi ansatz, emerged in description of hierarchical DNA folding in a chromosomes [20].

B. Ultrametricity and rare-event statistics

In the context of high dimensionality, randomness, and rare-event statistics, ultrametricity emerges in diverse subjects, e.g., in genetic trees, in the behavior of hierarchically organized ecological, social, and economic systems, etc. The fact that ultrametricity is naturally rooted in high dimensionality, randomness, and sparse statistics, has been unambiguously shown recently in [21]. It was proven that in a D -dimensional Euclidean space the distances between points in a highly sparse sampling tends to the ultrametric distances as $D \rightarrow \infty$.

In this work we bring another intriguing connection between rare-event statistics and ultrametricity. We demonstrate that ultrametricity is inherit for ensembles of sparse graphs. In particular, we compare a limiting behavior of spectral density of bi-diagonal random matrices with the behavior of the function $f(z) = \sqrt{-\log |\eta(z)|}$, where $\eta(z)$ is the Dedekind η -function and z tends to rational points on real axis, and conjecture that the eigenvalue distribution of sparse ensembles shares some number-theoretic properties related to ultrametricity.

Our finding provides an example of a nontrivial manifestation of rare-event Bernoullian noise in some collective observables in large systems (like spectral density, for example). From this point of view, our work should warn researchers working with effects of ultra-low doses of chemicals: in the statistical analysis of, say, absorption spectra of extremely diluted solutions of chain-like polymer molecules, the valuable signal should be clearly purified from the background noise, which itself could have a very peculiar ultrametric shape.

The paper is organized as follows. In the Section II we formulate the model under consideration and describe the results of our numerical simulations. In what follows we pay attention to statistics of linear subgraphs, presented by bi-diagonal matrices. We derive the distribution of the fraction of linear subgraphs at the percolation threshold in the Section III, and study the spectral properties of Bernoulli ensembles of bi-diagonal matrices in the Section IV. The number-theoretic structure of limiting spectral statistics in ensembles of random bi-diagonal matrices is discussed at length of the Section V. The results of the work and the relation with other natural systems shearing similar number-theoretic properties, are summarized in the Discussion.

II. THE MODEL

The Bernoulli matrix model is defined as follows. Take a large $N \times N$ symmetric matrix A . The matrix elements, $a_{ij} = a_{ji}$, are assumed to be independent, identically distributed, binary random variables, such that diagonal elements vanish, $a_{ii} = 0$, and

$$a_{ij} = \begin{cases} 1 & \text{with probability } q \\ 0 & \text{with probability } 1 - q \end{cases} \quad (1)$$

for $i \neq j$. The matrix A can be regarded as an adjacency matrix of a random Erdős-Rényi graph, \mathcal{G} , without self-connections. The eigenvalues, λ_n ($n = 1, \dots, N$), of the symmetric matrix A are real. Let $\rho(\lambda)$ be the eigenvalue density of the ensemble of such matrices. For $N \gg 1$ the limiting shape of $\rho(\lambda)$ is known in various cases. If $q = \mathcal{O}(1)$, the function $\rho(\lambda)$ tends to the Wigner semicircle law, $\sqrt{4N - \lambda^2}$, typical for the Gaussian matrix ensembles. For $q = c/N$ ($c > 1$), the matrix A is *sparse* and the density $\rho(\lambda)$ of the sparse matrix ensemble has singularities at finite values of c [22–24]. In Refs. [25–28], the behavior of the spectral density, $\rho(\lambda)$, has been analyzed near $c = 1$ ($N \gg 1$). It has been pointed out that the function $\rho(\lambda)$ becomes more and more singular as c approaches 1 from above.

In ensemble of matrices A the percolation point is $q_c = 1/N$ (the true phase transition is defined in the thermodynamic limit, $N \rightarrow \infty$, only). For $q > q_c$, the graph \mathcal{G} has a giant component whose size is proportional to N , accompanied with numerous small components (almost all of them are trees); for $q < q_c$, the graph \mathcal{G} is composed of the collection of small components (almost all of them are again trees). For $qN \ll 1$ the isolated vertices dominate, contributing to the trivial spectral density, $\rho(\lambda) = \delta(\lambda)$. The most interesting region is close to q_c . In the scaling window $|q - q_c| = \mathcal{O}(N^{-4/3})$, the size of the largest component scales as $N^{2/3}$. In what follows we discuss the behavior of the spectral density $\rho(\lambda)$ for q in this region.

An example of a random graph with a randomly generated $N \times N = 500 \times 500$ adjacency matrix at $q = 2.0028 \times 10^{-3}$ is shown in Fig. 1. Note that the components are predominantly linear subgraphs. In other realizations we observed components with a single loop, but they are rare, usually a few per realization; more complicated components, namely those with more than one loop, appear very rarely.

The whole spectrum of a symmetric matrix A in a sparse regime just below the percolation point consists of singular peaks coming from the graph components (maximal disjoint subgraphs constituting the random graph). For a subgraph of size k , the corresponding adjacency matrix A_k is of size $k \times k$, and the spectrum is determined by the equation $\det(A_k - \lambda I_k) = 0$, where I_k is the $k \times k$ identity matrix. There are k eigenvalues; apart from k , the eigenvalues depend on the topology of the subgraph. The heights of peaks are the eigenvalue multiplicities, the eigenvalue $\lambda = 0$ appears with the highest multiplicity. For instance, the graph shown in Fig. 1 has 222 subgraphs of size one (isolated vertices) each of them contributing to $\lambda = 0$. There are also 10 linear chains of length three, 3 chains of size five, and 1 chain of size seven contributing to $\lambda = 0$; generally any chain of odd length has eigenvalue 0. Trees which are not linear chains also can have eigenvalue 0. One can bipartite any tree into two disjoint parts X_1 and X_2 with edges only between X_1 and X_2 . If the sizes $n_1 = |X_1|$ and $n_2 = |X_2|$ of parts are different, say $n_1 > n_2$, the adjacency matrix has eigenvalue 0 with multiplicity at least $n_1 - n_2$.

The singular spectral density, $\rho(\lambda)$, for an ensemble of 1000 random symmetric matrices A , each of size 500×500 and generated with probability $q = 2.0028 \times 10^{-3}$, is shown in Fig. 2a. The spectrum possesses the regular ultrametric hierarchical structure which becomes even more profound while plotting $\ln \rho(\lambda)$, as shown in Fig. 2b. Near the value $\lambda = \pm 2$ the enveloping curve of the function $\ln \rho(\lambda)$ exhibits a fracture, clearly visible in the insert of Fig. 2a, where the spectrum is replotted as $\rho^{1/3}(\lambda)$ and the slopes are depicted by the dashed lines s_1 and s_2 . (The choice of the exponent $1/3$ for the curve $\rho^{1/3}(\lambda)$ is an "experimental" selection: we have just noted that such a scaling gives the best linear shape for the enveloping curve and clearly demonstrates the fracture of the shape near $\lambda = \pm 2$.)

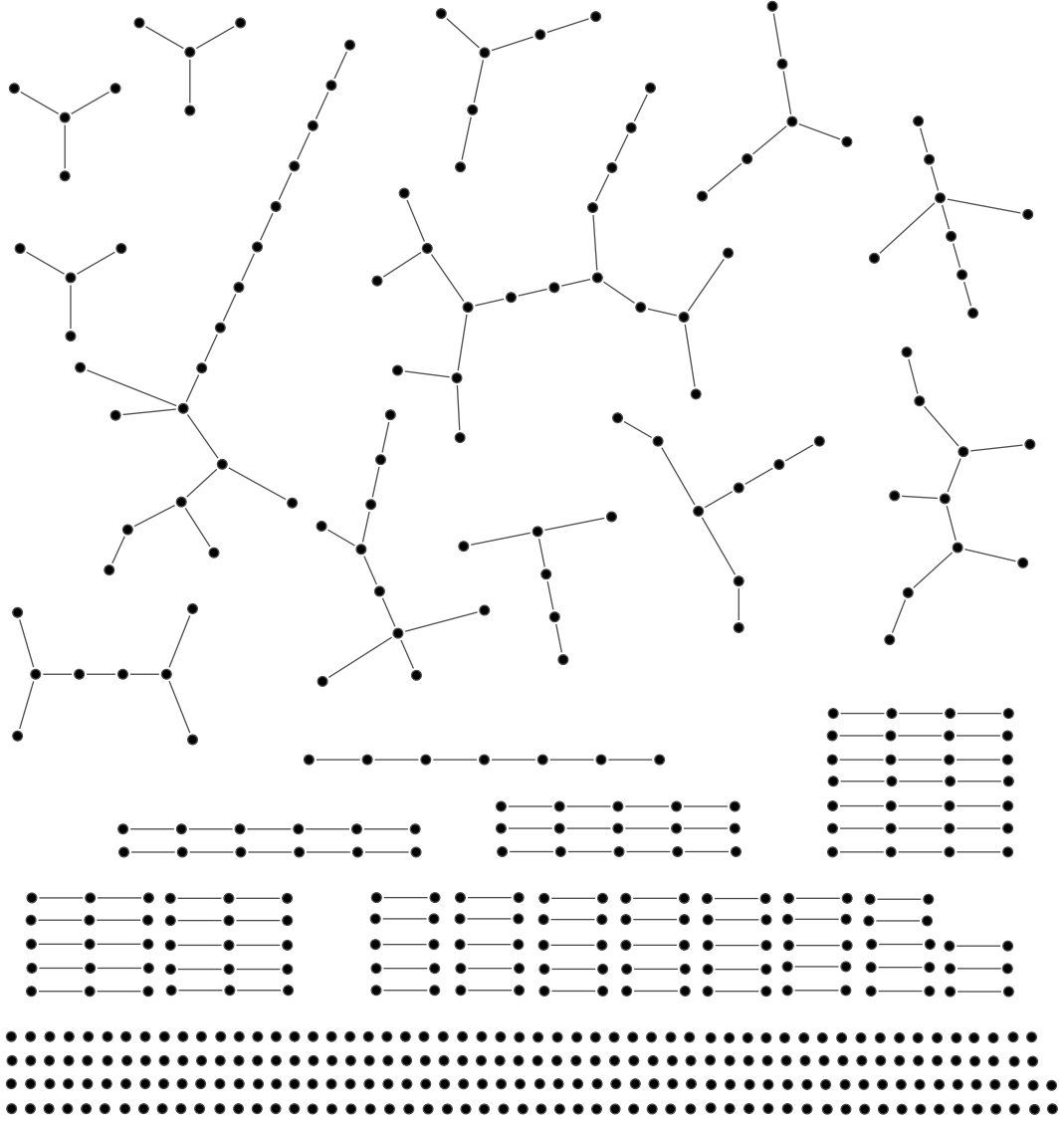


Figure 1: Typical sample of components constituting a random graph associated with an adjacency matrix A generated by the rule (1); the parameters are $N = 500$ and $q = 2.0028 \times 10^{-3}$.

The largest eigenvalue of a tree can be roughly estimated as $|\lambda_{\max}^{\text{tree}}| = 2\sqrt{p-1}$, where p is the maximal vertex degree of a tree (see e.g. [29–32]). For linear chains, the maximal degree is $p = 2$ (when the chain length is at least three) and the maximal eigenvalue is $|\lambda_{\max}^{\text{lin}}| = 2$.

In the bulk of the spectrum of sparse ensemble, the randomly branching trees (with $p = 3$) and linear chains (with $p = 2$) dominate (some estimates are given in the following sections). For $|\lambda| > |\lambda_{\max}^{\text{lin}}| = 2$, the linear chains do not contribute to the spectral density, and only graphs with $p \geq 3$ remain. For Erdős-Rényi random graphs, the fraction of trees with $p \geq 4$ is finite but numerically small. With a high probability, there are no such trees for sufficiently small N . This explains the maximal eigenvalue $|\lambda_{\max}^{\text{tree}}| = 2\sqrt{2} \approx 2.83$, coming from graphs with $p = 3$, which we observed in simulations (Fig. 2b).

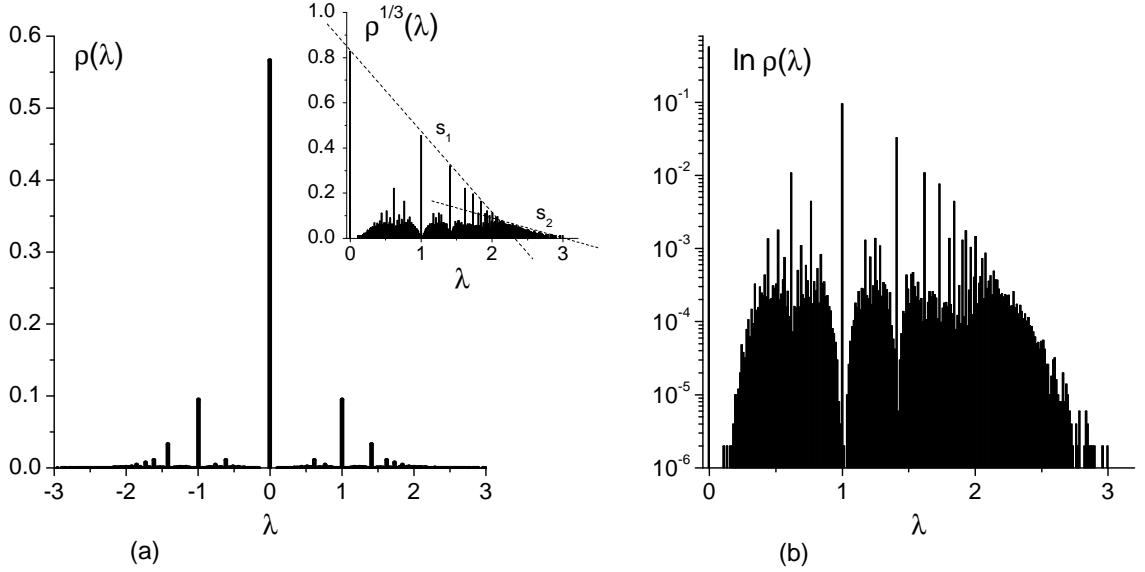


Figure 2: (a) Spectral density $\rho(\lambda)$ for ensemble of 1000 symmetric random 500×500 Bernoulli matrices, generated at $q = 2.0028 \times 10^{-3}$ possesses regular ultrametric structure. The fracture of the enveloping curve near $\lambda = \pm 2$ is clearly visible in the insert, where $\rho^{1/3}(\lambda)$ is plotted; (b) the plot of $\ln \rho(\lambda)$ shows the ultrametric structure of the spectral density and terminates close to the maximal eigenvalue for 3-branching trees, $|\lambda_{\max}^{\text{tree}}| = 2\sqrt{2} \approx 2.83$.

III. DISTRIBUTION OF LINEAR CHAINS AT THE PERCOLATION THRESHOLD

The spectral analysis of ensemble of random adjacency matrices A at the percolation threshold is cumbersome since the spectra of the components depend on their topology. For random Erdős-Rényi graphs, however, almost all components are trees (see e.g. [33, 34]). Furthermore, at the percolation point (and certainly below it) the majority of subgraphs are linear chains as we show below. The spectrum of each linear chain is easy to compute, so to get the spectral density, $\rho_{\text{lin}}(\lambda)$, of an ensemble of chain-like subgraphs we should know the distribution, Q_n , of linear chains of length n at the percolation point. Below we derive an explicit expression for Q_n . We first recall a kinetic theory approach which allows one to determine the size distribution of generic clusters (irrespective of their topology) and then adopt this framework to the computation of the size distribution of linear chains.

A. General formalism: Cluster size distribution

A cluster is a maximal connected component. To determine the cluster size distribution we employ a dynamical framework (see e.g. [34]), namely we begin with a disjoint graph represented by a collection of N separated vertices and start linking them with a constant rate. When two vertices from distinct clusters become linked, the clusters join.

Generally, there are $i \times j$ ways to join distinct clusters of size i and j . Hence, the clusters

undergo an aggregation process

$$(i, j) \xrightarrow{ij/(2N)} i + j \quad (2)$$

which proceeds with a rate proportional to the product of their sizes. We want to determine the cluster size distribution, $c_k(t)$. By definition, $c_k(t) = N_k(t)/N$, where $N_k(t)$ is the total number of components with k nodes. In (2) the linking rate is chosen to be $(2N)^{-1}$. With this choice, the cluster size distribution $c_k(t)$ satisfies

$$\frac{dc_k}{dt} = \frac{1}{2} \sum_{i+j=k} (ic_i)(jc_j) - k c_k \quad (3)$$

The initial condition is $c_k(0) = \delta_{k,1}$. The gain term in (3) accounts for clusters generated by joining two smaller clusters whose sizes sum up to k . The second term on the right-hand side of Eq. (3) represents loss due to linking of clusters of size k to other clusters. The corresponding gain and loss rates follow from the aggregation rule (2). Stopping the aggregation process at time t leads to a random graph with adjacency matrix characterized by the connectivity probability $q = t/N$.

Equations (3) have been solved via various techniques [34]. Here we recall one approach as we shall use it in the next sub-section to determine the distribution of linear chains. The idea is to eliminate the time dependence. Solving (3) recursively one gets

$$c_1(t) = e^{-t}; \quad c_2(t) = \frac{1}{2} t e^{-2t}; \quad c_3(t) = \frac{1}{2} t^2 e^{-3t}; \quad c_4(t) = \frac{2}{3} t^3 e^{-4t}; \quad \text{etc} \quad (4)$$

This suggests to seek the solution in the form

$$c_k(t) = C_k t^{k-1} e^{-kt} \quad (5)$$

Substituting this form into (3), we find that the coefficients C_k can be obtained recursively starting from $C_1 = 1$ and using

$$(k-1) C_k = \frac{1}{2} \sum_{i+j=k} (iC_i)(jC_j) \quad (6)$$

for $k > 1$. Using a generating function technique [35, 36] one gets $C_k = \frac{k^{k-2}}{k!}$. (The k^{k-2} factor is the Cayley formula [36] for the number of labelled trees on k vertices; Eq. (7) can be derived via a direct combinatorial computation.) Thus

$$c_k(t) = \frac{k^{k-2}}{k!} t^{k-1} e^{-kt} \quad (7)$$

The large k tail is always exponential apart from the percolation point $t = t_c = 1$ where

$$c_k(1) \simeq (2\pi)^{-1/2} k^{-5/2} \quad \text{when} \quad k \gg 1 \quad (8)$$

B. Distribution of linear chains

Consider now linear clusters (open chains). Let $Q_k(t)$ be the density of linear clusters of length $k \geq 2$. We have $Q_2 = c_2$. For $k \geq 3$ the chain size distribution evolves according to

$$\frac{dQ_k}{dt} = 2c_1 Q_{k-1} + 2 \sum_{i+j=k} Q_i Q_j - k Q_k \quad (9)$$

where the summation is taken over $i \geq 2, j \geq 2$. It is convenient to define $Q_1 = \frac{1}{2} c_1$. This allows us to rewrite (9) as

$$\frac{dQ_k}{dt} = 2 \sum_{i+j=k} Q_i Q_j - k Q_k \quad (10)$$

here the summation runs over all i, j satisfying $i + j = k$. Another virtue of the agreement $Q_1 = \frac{1}{2} c_1$ is that (10) is valid for all $k \geq 2$. Computing the first few densities we get

$$Q_1 = \frac{1}{2} c_1 = \frac{1}{2} e^{-t}; \quad Q_2 = c_2 = \frac{1}{2} t e^{-2t}; \quad Q_3 = c_3 = \frac{1}{2} t^2 e^{-3t}; \quad Q_4 = \frac{1}{2} t^3 e^{-4t} \quad (11)$$

etc (we have used (10) to determine Q_4). One is led to the conjectural behavior

$$Q_k(t) = \frac{1}{2} t^{k-1} e^{-kt} \quad (12)$$

This indeed provides the solution to (10) and the initial condition $Q_k(0) = \frac{1}{2} \delta_{k,1}$.

Thus, the linear chain length distribution (12) is *purely* exponential. Rewriting (12) as

$$Q_k(t) = \frac{1}{2t} e^{-k(t-\ln t)} \quad (13)$$

we see that the least-steepest decay occurs at the percolation point $t = t_c = 1$ where

$$Q_k \equiv Q_k(1) = \frac{1}{2} e^{-k} \quad (14)$$

The total *chain density* at the percolation point is

$$c_1 + \sum_{j \geq 2} Q_j(1) = \frac{1}{2} \frac{2 - e^{-1}}{e - 1} = 0.474928 \dots \quad (15)$$

Comparing this result with the total cluster density at the percolation point, $\sum_{k \geq 1} c_k(1) = \frac{1}{2}$, we conclude that at the percolation point almost 95% of clusters are linear chains. This justifies the abundance of linear chains (see Fig. 1).

Note that the total *mass density* of linear chains at the percolation point,

$$c_1 + \sum_{j \geq 2} j Q_j(1) = \frac{1}{2e} [1 + (1 - e^{-1})^{-2}] = 0.6442765 \dots, \quad (16)$$

is significantly smaller than unity. Large clusters are usually not linear and hence despite of their rareness they provide a notable contribution to the mass density.

C. Comments on more complicated components and different random graph models

Below the percolation point, the total number of components which are not trees remains finite in the thermodynamic limit $N \rightarrow \infty$, so it greatly fluctuates from realization to realization. The components which are not trees are *unicyclic*: each such component contains one loop. At the percolation point the total number of unicyclic components diverges but

very slowly, viz. as $\frac{1}{6} \ln N$. These results were established using probabilistic techniques, see [33]; they are also easily derivable using above dynamical approach (see [37]). Overall we see that the contribution of the non-trees is negligible.

The influence of the "shape" of the tree on the spectra is more notable. We have shown, however, that the majority of the clusters are not merely trees, but linear chains. In principle, one can modify the definition (1) to ensure that all components are chains. If each row (and column) of the adjacency matrix A has no more than two non-zero elements, $\sum_{j=1}^N a_{ij} \leq 2$ for $i = 1, \dots, N$, the corresponding graph is obviously the collection of linear chains (including loops) of various lengths.

Equivalently, the chain model can be defined dynamically by using the same procedure as before and checking that the constraint is obeyed. In other words, we attempt to link vertices randomly, but we actually draw a link only if the degrees of both vertices are smaller than 2. The cluster size distribution characterizing the chain model evolves according to master equation

$$\frac{dc_k}{dt} = \frac{1}{2} \sum_{i+j=k} K_{ij} c_i c_j - c_k \sum_{j \geq 1} K_{kj} c_j \quad (17)$$

with reaction rate matrix

$$K_{ij} = \begin{pmatrix} 1 & 2 & 2 & 2 & \cdots \\ 2 & 4 & 4 & 4 & \\ 2 & 4 & 4 & 4 & \\ 2 & 4 & 4 & 4 & \\ \vdots & & & & \ddots \end{pmatrix} \quad (18)$$

The initial condition is $c_k(0) = \delta_{k,1}$. A more general model with reaction rate matrix of the type (18) composed of three arbitrary rates has been studied in [38].

Using (17)–(18) one finds that the density of isolated vertices $c_1(t)$ and the total cluster density $c(t) = \sum_{j \geq 1} c_j(t)$ satisfy a pair of coupled rate equations:

$$\frac{dc_1}{dt} = -c_1(2c - c_1), \quad \frac{dc}{dt} = -\frac{1}{2}(2c - c_1)^2 \quad (19)$$

The initial condition becomes $c(0) = c_1(0) = 1$. Treating c as a function of c_1 we recast (19) into $\frac{dc}{dc_1} = \frac{c}{c_1} - \frac{1}{2}$, from which

$$c = c_1 \left[1 - \frac{1}{2} \ln c_1 \right] \quad (20)$$

Plugging this into the first equation (19) we obtain an implicit solution

$$\int_{c_1}^1 \frac{du}{u^2(1 - \ln u)} = t \quad (21)$$

The following densities can be found by re-writing the rate equations (17) in a manifestly

recursive form,

$$\begin{aligned}\frac{dc_2}{dt} &= \frac{1}{2}c_1^2 - 2c_2(2c - c_1), \\ \frac{dc_3}{dt} &= 2c_1c_2 - 2c_3(2c - c_1), \\ \frac{dc_4}{dt} &= 2c_1c_3 + 2c_2^2 - 2c_4(2c - c_1), \\ \frac{dc_5}{dt} &= 2c_1c_4 + 4c_2c_3 - 2c_5(2c - c_1),\end{aligned}$$

etc., and solving them recursively.

We use the notation c_k , but we emphasize that clusters generated by the above procedure are linear chains. There is no percolation transition in this model, so when should we stop the process? Recall that for Erdős-Rényi random graphs, the percolation occurs when the average degree is equal to one. In the model (17)–(18), the fraction of vertices of degree zero (d_0), vertices of degree one (d_1), and vertices of degree two (d_2) are

$$\begin{aligned}d_0 &= c_1 \\ d_1 &= 2 \sum_{j \geq 2} c_j = 2(c - c_1) \\ d_2 &= 1 - d_0 - d_1 = 1 - c_1 - 2(c - c_1)\end{aligned}$$

The average degree $\langle d \rangle = d_1 + 2d_2$ is therefore $\langle d \rangle = 2(1 - c)$. Thus, in analogy with Erdős-Rényi random graphs, we stop the process when the average degree reaches unity, the corresponding time t_s is found from $c(t_s) = \frac{1}{2}$. Another choice is to run the process till the very end when the system turns into a collection of closed loops, see [39] for this and other models of ring formation.

One can generalize above model by requiring that each row (and column) of the adjacency matrix A has no more than p non-zero elements: $\sum_{j=1}^N a_{ij} \leq p$ for $i = 1, \dots, N$. This model can be investigated in the dynamic framework. Unfortunately, it has proved to be much more difficult for the analysis [40] than the extreme versions corresponding to the chain model ($p = 2$) and the Erdős-Rényi variant ($p = N - 1$).

IV. SPECTRAL DENSITY OF CANONICAL ENSEMBLE OF BI-DIAGONAL RANDOM MATRICES

A. General formalism: Spectral density

We have seen in the Section III that the fraction of linear clusters of various lengths is about 95% at the percolation threshold in the ensemble of all sparse graphs. This justifies our study of ensembles of linear clusters (open chain and loops) only.

It is convenient to represent the linear cluster ensemble at the percolation point (with the distribution Q_k in chain lengths) by the bi-diagonal $N \times N$ symmetric matrix $B = \{b_{ij}\}$,

where

$$B = \begin{pmatrix} 0 & x_1 & 0 & 0 & \cdots \\ x_1 & 0 & x_2 & 0 & \\ 0 & x_2 & 0 & x_3 & \\ 0 & 0 & x_3 & 0 & \\ \vdots & & & & \ddots \end{pmatrix} \quad (22)$$

and the distribution of each $b_{i,i+1} = b_{i+1,i} = x_i$ ($i = 1, \dots, N$) is Bernoullian:

$$x_i = \begin{cases} 1 & \text{with probability } q \\ 0 & \text{with probability } 1 - q \end{cases} \quad (23)$$

Due to ergodicity, the spectral density, $\rho_{\text{lin}}(\lambda)$, of the ensemble of $N \times N$ matrices B in the limit $N \rightarrow \infty$ coincides with the spectral density of an individual matrix B and can be computed straightforwardly. Note that for any $x_k = 0$, the matrix B splits into independent fragments. So, it is instructive to consider subchains consisting of sequential set of $x_k = 1$. The probability to have a subchain with n consecutive 1 is $\tilde{Q}_n = q^n$, as it follows from (23). Comparing \tilde{Q}_n with the distribution $Q_n = \frac{1}{2}e^{-n}$ of linear subchains in ultra-sparse matrix ensemble at the percolation point (see (14)), we conclude that $q = e^{-1}$.

In what follows we derive $\rho_{\text{lin}}(\lambda)$ for arbitrary values of q . This allows us to investigate the spectral density of the ensemble of bi-diagonal random operators in the limit $q \rightarrow 1$ and uncover its number-theoretic structure.

The spectrum of linear chain of length n (i.e. of symmetric $n \times n$ bi-diagonal matrix B with all $x_k = 1$, $k = 1, \dots, n$) reads

$$\lambda_k = 2 \cos \frac{\pi k}{n+1}; \quad (k = 1, \dots, n) \quad (24)$$

The spectral density of the ensemble of $N \times N$ random matrices B with the Poissonian distribution of matrix elements can be written as

$$\begin{aligned} \rho_{\text{lin}}(\lambda) &= \lim_{N \rightarrow \infty} \frac{1}{N} \left\langle \sum_{k=1}^N \delta(\lambda - \lambda_k) \right\rangle \\ &= \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0}} \frac{\varepsilon}{\pi N} \text{Im} \langle G_N(\lambda - i\varepsilon) \rangle \\ &= \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0}} \frac{\varepsilon}{\pi N} \sum_{n=1}^N \tilde{Q}_n \text{Im} G_n(\lambda - i\varepsilon) \end{aligned} \quad (25)$$

where we have used an identity

$$\lim_{\varepsilon \rightarrow +0} \text{Im} \frac{1}{\lambda \pm i\varepsilon} = \mp i\pi \delta(\lambda)$$

The Green function

$$G_n(\lambda - i\varepsilon) = \sum_{k=1}^n \frac{1}{\lambda - \lambda_k - i\varepsilon} \quad (26)$$

with λ_k given by (24), is associated with each particular gapless matrix B of n sequential ones on the sub-diagonals, while $\langle \dots \rangle$ means averaging over the distribution $\tilde{Q}_n = q^n$.

Collecting (24), (25) and (26), we find an explicit expression for the density of states, $\rho_{\text{lin}}(\lambda)$:

$$\rho_{\text{lin}}(\lambda) = \lim_{\substack{N \rightarrow \infty \\ \varepsilon \rightarrow 0}} \frac{\varepsilon}{\pi N} \sum_{n=1}^N q^n \sum_{k=1}^n \frac{1}{\left(\lambda - 2 \cos \frac{\pi k}{n+1}\right)^2 + \varepsilon^2} \quad (27)$$

The sample plots $\rho_{\text{lin}}(\lambda)$ for two different values of q , namely for $q = 0.7$ and $q = e^{-1} \approx 0.37$ computed numerically via (27) with $\varepsilon = 3 \times 10^{-3}$ are shown in the Fig. 3 for $N = 10^3$.

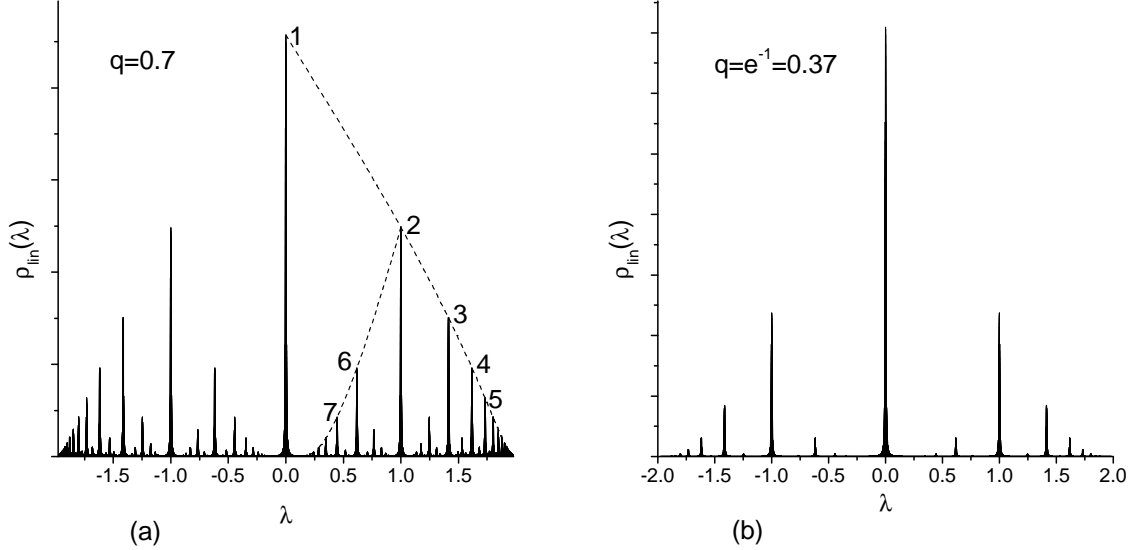


Figure 3: The spectral density $\rho_{\text{lin}}(\lambda)$ for the ensemble of bi-diagonal matrices of size $N = 10^3$ at $q = 0.7$ (a) and $q = e^{-1} \approx 0.37$ (b); the regularization parameter ε is taken $\varepsilon = 2 \times 10^{-3}$.

B. Analysis of enveloping curves and tails of spectral density $\rho_{\text{lin}}(\lambda)$

One can compute the enveloping curves for any monotonic sequence of peaks. In Fig. 3, we show two series of sequential peaks: $S_1 = \{1-2-3-4-5-\dots\}$ and $S_2 = \{2-6-7-\dots\}$. Any monotonic sequence of peaks corresponds to the set of eigenvalues λ_k constructed on the basis of Farey sequence [41]. For example, as shown below, the peaks in the series S_1 are located at $\lambda_k = -2 \cos \frac{\pi k}{k+1}$, ($k = 1, 2, \dots$), while the peaks in the series S_2 are located at $\lambda_{k'} = -2 \cos \frac{\pi k'}{2k'-1}$, ($k' = 2, 3, \dots$). Positions of peaks obey the following rule: let $\{\lambda_{k-1}, \lambda_k, \lambda_{k+1}\}$ be three consecutive monotonically ordered peaks (e.g., peaks 2-3-4 in Fig. 3), and let

$$\lambda_{k-1} = -2 \cos \frac{\pi p_{k-1}}{q_{k-1}}, \quad \lambda_{k+1} = -2 \cos \frac{\pi p_{k+1}}{q_{k+1}}$$

where p_k and q_k ($k = 1, \dots, N$) are coprimes. The position of the intermediate peak, λ_k , is

$$\lambda_k = -2 \cos \frac{\pi p_k}{q_k}, \quad \frac{p_k}{q_k} = \frac{p_{k-1}}{q_{k-1}} \oplus \frac{p_{k+1}}{q_{k+1}} \equiv \frac{p_{k-1} + p_{k+1}}{q_{k-1} + q_{k+1}} \quad (28)$$

The sequences of coprime fractions constructed via the \oplus addition are known as Farey sequences. A simple geometric model behind the Farey sequence, known as Ford circles [42], is shown in Fig. 4a. Take the segment $[0, 1]$ and draw two circles O_1 and O_2 both of radius $r = \frac{1}{2}$, which touch each other, and the segment at the points 0 and 1. Now inscribe a new circle O_3 touching O_1 , O_2 and $[0, 1]$. Where is the position of the new circle along the segment? The generic recursive algorithm is shown in Fig. 4b and constitutes the Farey sequence construction.

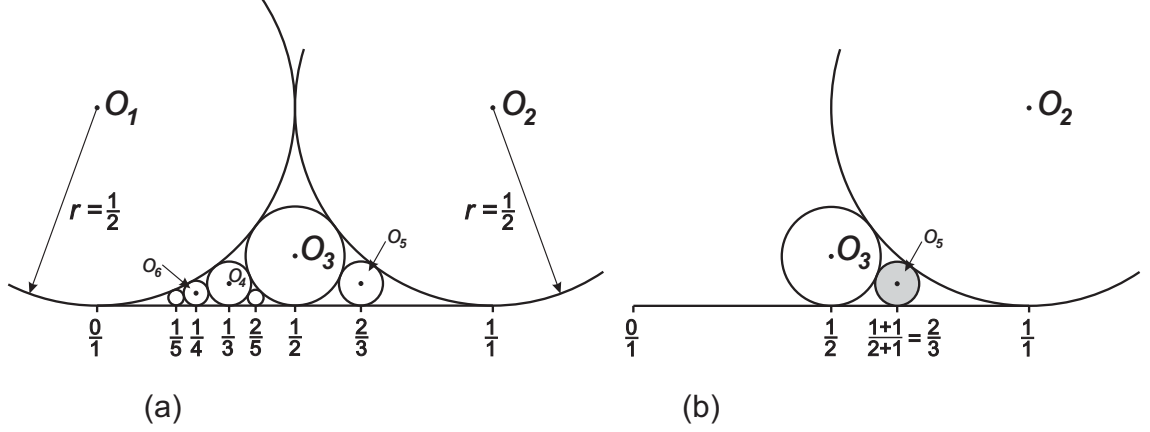


Figure 4: Ford circles as illustration of the Farey sequence construction: a) Each circle touches two neighbors (right and left) and the segment; b) Determination of the position of newly generated circle via the \oplus addition: $\frac{p_{k-1}}{q_{k-1}} \oplus \frac{p_{k+1}}{q_{k+1}} = \frac{p_{k-1}+p_{k+1}}{q_{k-1}+q_{k+1}}$.

Consider two examples related to the set of eigenvalues.

- The monotonic series S_1 is can be constructed recursively by taking

$$\lambda_1 \equiv -2 \cos \frac{\pi}{2} = 0, \quad \lambda_\infty \equiv -2 \cos \frac{\pi}{1} = 2$$

The next-to-highest peak 2 is located at

$$\lambda_2 = -2 \cos \left[\pi \left(\frac{1}{2} \oplus \frac{1}{1} \right) \right] = -2 \cos \frac{2\pi}{3} = 1$$

Peak 3 is located at

$$\lambda_3 = -2 \cos \left[\pi \left(\frac{2}{3} \oplus \frac{1}{1} \right) \right] = -2 \cos \frac{3\pi}{4} = \sqrt{2}$$

Peak 4 is located at

$$\lambda_4 = -2 \cos \left[\pi \left(\frac{3}{4} \oplus \frac{1}{1} \right) \right] = -2 \cos \frac{4\pi}{5} = \frac{\sqrt{5}+1}{2}$$

etc. The generic expression for λ_k in the series S_1 reads

$$\lambda_k = -2 \cos \frac{k\pi}{k+1}, \quad (k = 1, 2, \dots) \quad (29)$$

- The monotonic series S_2 is can be constructed similarly by taking as reference points λ_1 and λ_2 . For peak 6 we have

$$\lambda_6 = -2 \cos \left[\pi \left(\frac{1}{2} \oplus \frac{2}{3} \right) \right] = -2 \cos \frac{3\pi}{5} = \frac{\sqrt{5}-1}{2}$$

Peak 7 is located at

$$\lambda_7 = -2 \cos \left[\pi \left(\frac{1}{2} \oplus \frac{3}{5} \right) \right] = -2 \cos \frac{4\pi}{7}$$

etc. The generic expression for $\lambda_{k'}$ in the series S_2 reads

$$\lambda_{k'} = -2 \cos \frac{k'\pi}{2k'-1}, \quad (k' = 2, 3, \dots) \quad (30)$$

As follows from (27), the function $\rho_{\text{lin}}(\lambda)$ is nonzero only at λ_k . In Fig. 5 we show the set of eigenvalues $\lambda_n(k) = -2 \cos \frac{\pi k}{n+1}$ for $n = 1, \dots, 50$. Every point designates some eigenvalue, λ_k (see (24)); the points along each curve correspond to *different* values k for a *specific* value n . The collection of eigenvalues in each horizontal line, coming from different n , gives the multiplicity of corresponding eigenvalue λ_k and contributes to the height of a peak in the spectral density $\rho_{\text{lin}}(\lambda)$.

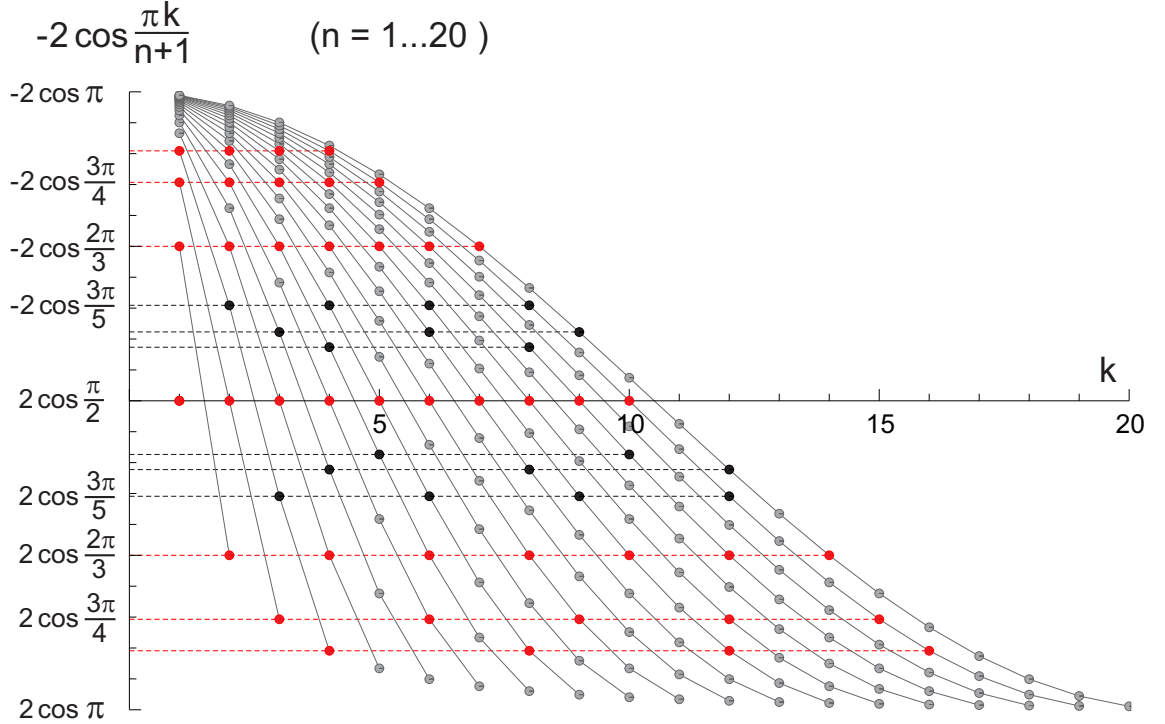


Figure 5: Family of 20 curves $w_n(k) = -2 \cos \frac{\pi k}{n+1}$. Each curve corresponds to a particular $n = 1, \dots, 20$ from left to the right, the points along the curves mark values $k = 1, \dots, n$. Each horizontal dashed line is the multiplicity of the corresponding root and contributes to the height of the peak in the spectral density.

The heights of peaks in any monotonic sequence (like S_1 or S_2) can be computed directly from (27) by summing geometric series along each horizontal line in Fig. 5. For example, for the sequences S_1 and S_2 we have:

$$S_1 : \begin{cases} \lambda_1 = -2 \cos \frac{\pi}{2} & \rho^{S_1}(\lambda_1) = q^1 + q^3 + q^5 + q^9 + \dots \\ \lambda_2 = -2 \cos \frac{2\pi}{3} & \rho^{S_1}(\lambda_2) = q^2 + q^5 + q^8 + q^{11} + \dots \\ \dots & \dots \\ \lambda_k = -2 \cos \frac{\pi k}{k+1} & \rho^{S_1}(\lambda_k) = \sum_{s=1}^{\infty} q^{(k+1)s-1} = \frac{q^k}{1-q^{k+1}} \quad (k = 1, 2, \dots) \end{cases} \quad (31)$$

and

$$S_2 : \begin{cases} \lambda_2 = -2 \cos \frac{2\pi}{3} & \rho^{S_2}(\lambda_2) = q^2 + q^5 + q^8 + q^{11} + \dots \\ \lambda_3 = -2 \cos \frac{3\pi}{5} & \rho^{S_2}(\lambda_3) = q^4 + q^9 + q^{14} + q^{19} + \dots \\ \dots & \dots \\ \lambda_{k'} = -2 \cos \frac{\pi k'}{2k'-1} & \rho^{S_2}(\lambda_{k'}) = \sum_{s=1}^{\infty} q^{(2k'-1)s-1} = \frac{q^{2k'-2}}{1-q^{2k'-1}} \quad (k = 2, 3, \dots) \end{cases} \quad (32)$$

Note that the expressions (31)–(32) are tightly linked to so-called visibility diagram (our notion is slightly modified with respect to a visibility diagram defined in [43]). Consider the square lattice of integer points (m, n) and add a weight q^m to each vertical row. Emit rays at rational angles $\alpha_{p,q} = \arctan \frac{p}{q}$ from the point $(0, 0)$ within the wedge $[\pi/4, \pi/2]$ in the positive direction as shown in Fig. 6.

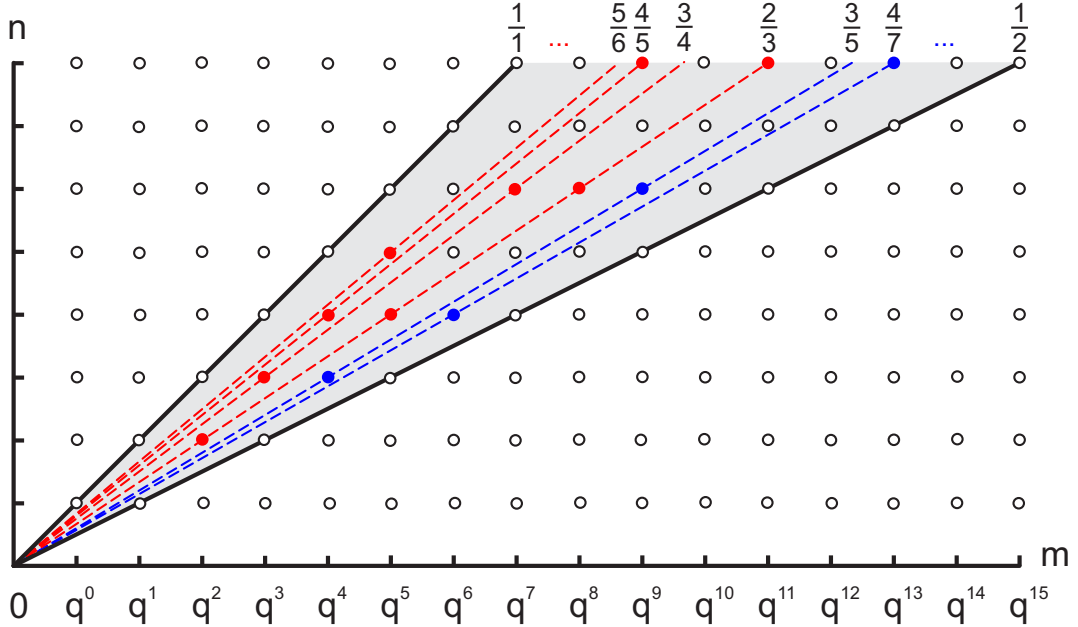


Figure 6: Visibility diagram. Each point in the vertical row m carries a weight q^m . Integer points within the wedge $[\pi/4, \pi/2]$ are designated by circles. The dashed rays are drawn at rational angles $\alpha_{p,q} = \arctan \frac{p}{q}$. The weights q^m corresponding to marked integer points, are summed up along the rays.

Let us sum up the weights, q^m , of integer points along each ray emitted at the rational slope $\frac{p}{q}$. For example, consider the sequence of slopes $\tan \alpha_{1,k} = \frac{1}{k}$ with $k = 1, 2, \dots$ and $\tan \alpha_{k',2k'-1} = \frac{k'}{2k'-1}$ shown in Fig. 6. One can verify that these geometric series exactly reproduce the sequences S_1 and S_2 in Eqs. (31)–(32) at corresponding k and k' .

Two parametrically defined enveloping curves, namely $(\lambda_k, \rho^{S_1}(k))$ for the sequence S_1 ($k = 1, 2, \dots$) and $(\lambda_{k'}, \rho^{S_2}(k'))$ for the sequence S_2 ($k' = 2, 3, \dots$), are shown in Fig. 3a by dashed lines. Equations (31)–(32) allow to find the singular tails of distributions $\rho^{S_1}(\lambda)$ as $\lambda \rightarrow 2$ and $\rho^{S_2}(\lambda)$ as $\lambda \rightarrow 0$. Expand now $\rho^{S_1}(\lambda_k)$ in (31) at $k \rightarrow \infty$ and $\rho^{S_2}(\lambda_{k'})$ at $k' \rightarrow \infty$:

$$\begin{aligned} S_1 : \lambda &= -2 \cos \frac{\pi k}{k+1} \Big|_{k \rightarrow \infty} \rightarrow 2 - \frac{\pi^2}{k^2}, & \rho^{S_1}(\lambda_k) \Big|_{k \rightarrow \infty} &\rightarrow q^k \\ S_2 : \lambda' &= -2 \cos \frac{\pi k'}{2k'-1} \Big|_{k' \rightarrow \infty} \rightarrow \frac{\pi}{2k'}, & \rho^{S_2}(\lambda_{k'}) \Big|_{k' \rightarrow \infty} &\rightarrow q^{2k'} \end{aligned} \quad (33)$$

Expressing k and k' in (33) in terms of λ and λ' , we get

$$k \rightarrow \frac{\pi}{\sqrt{2-\lambda}}, \quad k' \rightarrow \frac{\pi}{2\lambda'} \quad (34)$$

Asymptotics (34) together with (33) permit us to find the singular tails of the eigenvalue distribution $\rho_{\text{lin}}(\lambda)$ at $\lambda \rightarrow 2$ and $\lambda' \rightarrow 0$:

$$\rho_{\text{lin}}^{S_1}(\lambda \rightarrow 2) \rightarrow q^{\pi/\sqrt{2-\lambda}}, \quad \rho_{\text{lin}}^{S_2}(\lambda' \rightarrow 0) \rightarrow q^{\pi/\lambda'} \quad (35)$$

The edge singularity $\rho_{\text{lin}}^{S_1}(\lambda \rightarrow 2)$ at $\lambda \rightarrow 2$ in Eq. (35) of the eigenvalue distribution reproduces the corresponding Lifshitz tail of the one-dimensional Anderson localization in random Schrödinger operator [44, 45]:

$$\rho(E) \sim e^{-1/E^{d/2}}$$

where $E = 2 - \lambda$ and $d = 1$. The appearance of the edge singularity $\rho(E) \sim e^{-1/\sqrt{E}}$ in our situation is purely geometric, it is not relied on any entropy-energy-balance consideration like optimal fluctuation [44]. Moreover, we would like to emphasize some universality of the Lifshitz tail: typically, Anderson localization of type (35) appears for random three-diagonal operators with the randomness on the main diagonal, while in our case we have a randomness in off-diagonal symmetric "hopping rate".

V. SPECTRAL DENSITY OF BI-DIAGONAL MATRICES IN THE LIMIT $q \rightarrow 1$ AND THE DEDEKIND η -FUNCTION

In this Section we present arguments demonstrating that the singular spectral density, $\rho(\lambda)$, of ensemble of random bi-diagonal operators B defined in (22)–(23), in the asymptotic regime $q \rightarrow 1$ is connected to the limiting behavior of some well-known modular function, namely of Dedekind η -function. Our conclusion is based on observation that $\rho(\lambda)$ at singularities can be expressed in terms of Dedekind η -function at rational points near the real axis. Since our derivation is far from being rigorous, we support it by numeric computations, which demonstrate striking similarity of $\rho(\lambda)$ in the limit $q \rightarrow 1$ and a profile constructed on the basis of $\eta(z)$ for $\text{Im } z \rightarrow 0$.

A. Tails of bi-diagonal matrices at $q \rightarrow 1$

Generic expressions (31)–(32) allow to analyze the tails of the distribution $\rho_{\text{lin}}(\lambda)$ when $q \rightarrow 1$ and k is fixed. Namely, taking the limit $q \rightarrow 1$ in (31)–(32), we arrive at:

$$\begin{aligned}\rho_1^{S_1}(\lambda_k)|_{q \rightarrow 1} &= \frac{q^k}{1 - q^{k+1}} \Big|_{q \rightarrow 1} \rightarrow \frac{1}{(k+1)(1-q)} \\ \rho_1^{S_2}(\lambda_{k'})|_{q \rightarrow 1} &= \frac{q^{2k'-2}}{1 - q^{2k'-1}} \Big|_{q \rightarrow 1} \rightarrow \frac{1}{(2k'-1)(1-q)}\end{aligned}\quad (36)$$

These expressions lead us to the following asymptotics of tails of the distribution $\rho(\lambda)$ at $\lambda \approx 2 - \frac{\pi^2}{k^2} \Big|_{k \gg 1} \rightarrow 2$ and $\lambda' \approx \frac{\pi}{2k'} \Big|_{k \gg 1} \rightarrow 0$:

$$\begin{aligned}\rho_1^{S_1}|_{q \rightarrow 1}(\lambda \rightarrow 2) &\rightarrow \frac{\sqrt{2-\lambda}}{\pi(1-q)} \\ \rho_1^{S_2}|_{q \rightarrow 1}(\lambda' \rightarrow 0) &\rightarrow \frac{\lambda'}{\pi(1-q)}\end{aligned}\quad (37)$$

Similarly, any monotonic sequence, can be analyzed, and corresponding degeneracy of eigenvalue λ (i.e. height of every peak) can be obtained. The Fig. 6 makes this statement transparent: we sum lattice points with proper weights along rays emitted at rational angles from the origin. We see that all tails of intermediate monotonic sequences have scaling behavior as in the sequence $\rho_1^{S_2}(\lambda)$. Note also an essential difference of (37) with (35).

B. Limiting properties of the Dedekind η -function via duality relation

Let us demonstrate that the tails (37) appear in asymptotic analysis related to the Dedekind η -function. Recall the definition of the Dedekind η -function [46]:

$$\eta(z) = e^{\pi iz/12} \prod_{n=0}^{\infty} (1 - e^{2\pi i n z}) \quad (38)$$

The argument $z = x + iy$ is called modular parameter and $\eta(z)$ is defined for $\text{Im } z > 0$ only.

It is convenient to introduce the following "normalized" function

$$f(z) = C^{-1} |\eta(z)| (\text{Im } z)^{1/4}; \quad C = \left| \eta \left(\frac{1}{2} + i \frac{\sqrt{3}}{2} \right) \right| \left(\frac{\sqrt{3}}{2} \right)^{1/4} \approx 0.77230184... \quad (39)$$

The structure of $f(z)$ in the context of ultrametric landscape construction has been discussed in the paper [47]. At the same time, it is well known that Dedekind η -function shares properties related to hyperbolic geometry. To make the content of this paper selfconsistent, we summarize in Appendix facts, which demonstrate the connection between the hyperbolic and ultrametric structures of Dedekind η -function.

The function $f(z)$ satisfies the following duality relation

$$f\left(\left\{\frac{m}{k}\right\} + iy\right) = f\left(\left\{\frac{s}{k}\right\} + \frac{i}{k^2 y}\right); \quad ms - kr = 1; \quad \{m, s, k, r\} \in \mathbb{Z}; \quad y > 0 \quad (40)$$

where $\left\{\frac{m}{k}\right\}, \left\{\frac{s}{k}\right\}$ denote fractional parts of corresponding quotients. Two particular duality identities follow from (40):

$$\begin{aligned} f\left(\frac{k}{k+1} + iy\right) &= f\left(\frac{1}{k+1} + \frac{i}{(k+1)^2 y}\right) \\ f\left(\frac{k'}{2k'-1} + iy\right) &= f\left(\frac{2}{2k'-1} + \frac{i}{(2k'-1)^2 y}\right) \end{aligned} \quad (41)$$

Using (41) we can obtain the asymptotics of $\eta(z)$ at $y \rightarrow 0^+$. Take into account the relation of Dedekind η with Jacobi elliptic functions:

$$\vartheta_1'(0, e^{\pi iz}) = \eta^3(z) \quad (42)$$

where

$$\vartheta_1'(0, e^{\pi iz}) \equiv \frac{d\vartheta_1(u, e^{\pi iz})}{du} \Big|_{u=0} = e^{\pi iz/4} \sum_{n=0}^{\infty} (-1)^n (2n+1) e^{\pi in(n+1)z} \quad (43)$$

Rewrite (41) in terms of η -functions:

$$\left| \eta\left(\frac{k}{k+1} + iy\right) \right| = \frac{1}{\sqrt{(k+1)y}} \left| \eta\left(\frac{1}{k+1} + \frac{i}{(k+1)^2 y}\right) \right| \quad (44)$$

(Since the computations are identical for both equations in (41), we do not reproduce the derivation for the second one, and will write for it the final answer only). Applying (42)–(43) to (44), we get:

$$\left| \eta\left(\frac{k}{k+1} + iy\right) \right| = \frac{1}{\sqrt{(k+1)y}} \left| e^{\frac{\pi}{4}i\left(\frac{1}{k+1} + \frac{i}{(k+1)^2 y}\right)} \sum_{n=0}^{\infty} (-1)^n (2n+1) e^{\pi in(n+1)\left(\frac{1}{k+1} + \frac{i}{(k+1)^2 y}\right)} \right|^{1/3} \quad (45)$$

Equation (45) enables us to extract the leading asymptotics of $\left| \eta\left(\frac{k}{k+1} + iy\right) \right|$ in the $y \rightarrow 0^+$ limit. Noting that every term of the series in (45) for any $n \geq 1$ and for small y (i.e. for $y \ll (k+1)^{-2}$) converges exponentially fast, we have

$$\left| \eta\left(\frac{k}{k+1} + iy\right) \right| \Big|_{y \rightarrow 0^+} \rightarrow \frac{1}{\sqrt{(k+1)y}} e^{-\frac{\pi}{12(k+1)^2 y}}; \quad (46)$$

Similarly, we get the asymptotics for $\left| \eta\left(\frac{k'}{2k'-1} + iy\right) \right|$ at $y \rightarrow 0^+$ (valid for $y \ll (2k'-1)^{-2}$). Thus, we have finally:

$$\begin{aligned} \sqrt{-\ln \left| \eta\left(\frac{k}{k+1} + iy\right) \right|} \Big|_{y \rightarrow 0^+} &\rightarrow \frac{1}{k+1} \sqrt{\frac{\pi}{12y}} \\ \sqrt{-\ln \left| \eta\left(\frac{k'}{2k'-1} + iy\right) \right|} \Big|_{y \rightarrow 0^+} &\rightarrow \frac{1}{2k'-1} \sqrt{\frac{\pi}{12y}} \end{aligned} \quad (47)$$

In (47) we have dropped out the subleading terms at $y \rightarrow 0^+$, which are of order of $O(\ln y)$.

Substitute now for k and k' in (47) the asymptotic expansions $k(\lambda) \approx \frac{\pi}{\sqrt{2-\lambda}}$ and $k'(\lambda') \approx \frac{\pi}{\lambda}$ (see (34)). Since near the spectrum edges, $k \gg 1$ and $k' \gg 1$, we get

$$\sqrt{-\ln |\eta(\lambda + iy)|} \Big|_{\substack{\lambda \rightarrow 2 \\ y \rightarrow 0}} \rightarrow \frac{\sqrt{2-\lambda}}{\sqrt{12\pi y}} \quad \text{and} \quad \sqrt{-\ln |\eta(\lambda' + iy)|} \Big|_{\substack{\lambda' \rightarrow 0 \\ y \rightarrow 0}} \rightarrow \frac{\lambda'}{\sqrt{12\pi y}} \quad (48)$$

The tails in (48) exactly match the tails of the spectral density $\rho^A(\lambda)$ and $\rho^B(\lambda')$ in (37) if we make an identification $1 - q = \sqrt{12\pi y}$.

Repeating the same asymptotic analysis for any other monotonic sequence of peaks, corresponding to particular sequences of coprime fractions, we arrive at the singularity behavior (47). This is consistent with the conclusion that any monotonic sequence of peaks in a spectral density $\rho(\lambda)$ has the singular behavior as in Eq.(36).

Now we are led to the central conjecture of this work, which uncovers the number theoretic nature of the spectral density of exponential ensemble of large random bi-diagonal symmetric matrices.

Conjecture. Let $\rho_{\text{lin}}(\lambda, q)$ be the spectral density of ensemble of $N \times N$ bi-diagonal symmetric random matrices with the Bernoulli distribution of nonzero matrix elements defined in Eqs.(22)–(23). For $N \rightarrow \infty$ we have the relation:

$$\lim_{q \rightarrow 1^-} \frac{\rho_{\text{lin}}(\lambda, q)}{\sqrt{-\ln \left| \eta \left(\frac{1}{\pi} \arccos \left(-\frac{\lambda}{2} \right) + i \frac{(1-q)^2}{12\pi} \right) \right|}} = 1 \quad (49)$$

where η is the Dedekind η -function defined in (38). One can say, that Eq.(49) establishes the link between spectral density of random one-dimensional Schrödinger-type operators with the limiting values of modular function at rational points on the real axis.

Empirical evidence in favor of the above conjecture is presented in Fig. 7 where we plot three functions:

Fig. 7a: Numerically generated spectral density of ensemble of $N \times N$ ($N = 3000$) bi-diagonal Bernoulli random matrices B (see (22)) with the Poissonian distribution of “1” at $q = 0.98$;

Fig. 7b: Numerically computed eigenvalue density $\rho(\lambda, q)$ given by Eq.(27) for matrices of size $N = 3000$ and $q = 0.98$;

Fig. 7c: Analytic expression given by the right-hand side of (49) at $y = \frac{(1-q)^2}{12\pi} \approx 10^{-5}$.

VI. DISCUSSION

A. Ultrametricity and rare-event statistics

We investigated the spectral density of a sub-ensemble of sparse symmetric Bernoulli matrices near the percolation point. We showed that in the vicinity of the percolation

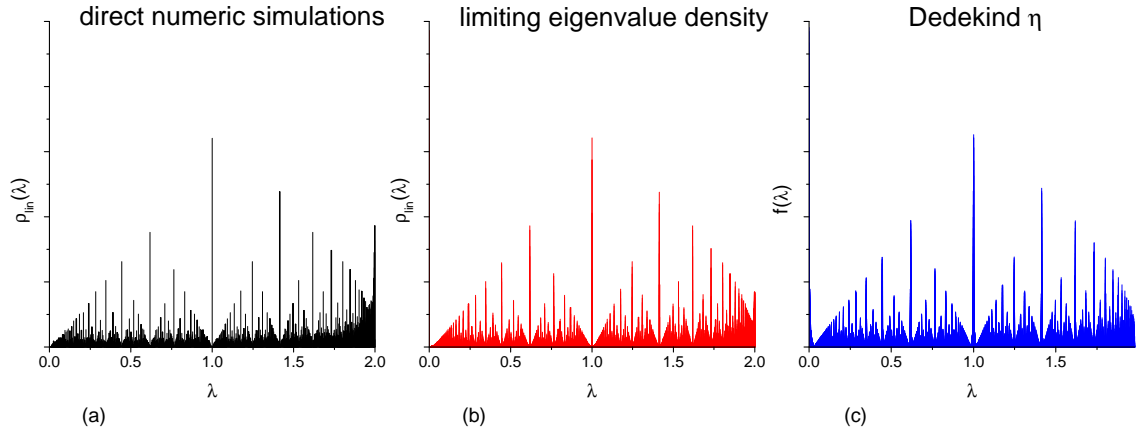


Figure 7: Spectral density $\rho_{\text{lin}}(\lambda)$ of ensemble of bi-diagonal Poissonian random matrices (shown only for $\lambda \geq 0$): a) direct numeric simulation; b) numeric visualization of Eq. (27); c) analytic construction provided by Eq. (49).

point linear chains constitute the overwhelming majority of clusters—about 95%. This led us to study the spectral density of ensemble of linear chains only. The advantage of this consideration is the possibility of exact treatment which allowed us to uncover the number-theoretic structure of the corresponding density of states.

In a more practical setting, our analysis is applicable to investigation of polydisperse solutions of linear polymers at any concentrations with Poissonian distribution in chain lengths, where the ultrametricity could be directly seen in experiments. Namely, we have checked that ultrametric structure of spectral density survives if one replaces the ensemble of Bernoullian two-diagonal adjacency matrices B defined by (22)–(23) by the ensemble of random Laplacian matrices [48]. Recall that the Laplacian matrix, $L = \{a_{ij}\}$, can be constructed from adjacency matrix, $B = \{b_{ij}\}$, as follows: $a_{ij} = -b_{ij}$ for $i \neq j$, and $a_{ii} = \sum_{j=1}^N b_{ij}$. Finding eigenvalues for Laplacian matrix L for linear chain, is equivalent to determining its vibration spectrum. Thus, having diluted solution of linear chains with exponential distribution in their lengths, and measuring the response on excitation of the system at various frequencies, we expect to see the signature of ultrametricity in adsorption spectra.

One tacit message coming from our work is that experimenting with biological activity of highly diluted solutions of biologically active substances, one should pay attention to a very peculiar structure of background noise originated from the rare-event statistics of dissolved clusters. In order to make conclusion about any biological activity of regarded chemical substance, the signal from background noise should be clearly identified. From this point of view, the work [49] seems very interesting, since it represents a rare example of careful attention to ultrametric-like distributions in biological and clinical data.

As concerns exact results, we showed that the spectral density, $\rho_{\text{lin}}(\lambda)$, in the ensemble of bi-diagonal random symmetric matrices with the Bernoulli distribution of matrix elements demonstrates a very peculiar ultrametric structure. Adding any constant value $b_{ii} = \bar{b}$ to the diagonal of the matrix B (see (22)–(23)) leads to the shift of the spectrum by the value $-\bar{b}$.

The derivation of the shape of the enveloping curve for the spectral density, $\rho_{\text{lin}}(\lambda)$, of

random Shrödinger-type operators presented by random symmetric bi-diagonal matrices, is based on simultaneous localization of eigenvalues, $\lambda_k = 2 \cos \frac{\pi}{k}$ ($k = 1, 2, \dots$), constituting the principal sequence, and of their multiplicativity (degeneracy), $\rho(\lambda_k) = \frac{q^k}{1-q^k}$ ($0 < q < 1$). Eliminating k from the parametric dependence $(\lambda_k, \rho(\lambda_k))$, one gets the signature of a Lifshitz tail for Anderson localization at the edge of the spectrum, $\rho(\lambda \rightarrow 2) \sim q^{-\pi/\sqrt{2-\lambda}}$, typical for the one-dimensional Schrödinger random operators. We emphasize that in our case the derivation is purely geometric.

Analyzing the limit $q \rightarrow 1$, we have demonstrated that the function $\rho_{\text{lin}}(\lambda)$ shares number-theoretic properties connected with the structure of the Dedekind η -function near the real axis at rational points. Intriguingly, the limiting behavior of the Dedekind function near real axis at rational points has tantalizing connections with Vassiliev knot invariants [50] and with quantum invariants of 3-manifolds [51]; a new concept of *quantum modular forms* was recently suggested [52] as a potential explanation.

B. Manifestation of the Dedekind η -function in natural science

Finally, we mention two other physical systems exhibiting a connection with the η function. In Ref. [53] it has been argued that buckling of a leaf of some plants (like lettuce or spinach) is related with isometric embedding of hyperbolic graphs (Cayley trees) into the 3D Euclidean space. The buckling is described by a function, $g(z)$, (Jacobian of transformation), which defines the deformation of the hyperbolic graph in its projection from the 3D Euclidean space into the complex plane (see also the Section 5.1 of [54]).

For a regular 3-branching Cayley tree, the function $g(z)$ can be expressed via the Dedekind η -function [47, 53]. It has been noted in [47] that properly normalized function $g(z)$, namely, $f(z) = C^{-1}|\eta(z)|(\text{Im } z)^{1/4}$, defined in Eq.(39), can be regarded as a continuous tree isometrically embedded into the half-plane $\text{Im } z > 0$ – see the Appendix for details. The cut of the function $f(z) \in [0.985, 1.0]$ is shown in Fig. 8a, while in Fig. 8b the same function is replotted in polar coordinates, obtained via the conformal mapping of the upper half-plane $\text{Im } z > 0$ to the unit disc. Note the striking similarity of Fig. 8a with the set of touching Ford circles (Fig. 4). In fact, the x -coordinates of centers of all lacunas in Fig. 8a obey the Farey construction.

Another well-known manifestation of number theory in natural science arises in the context of phyllotaxis [55]. The connection of the cell packing with the Farey sequences has been observed long time ago. It was not clear, however, why nature selects the Fibonacci sequence among other possible Farey sequences. A tantalizing answer to this question has been given by L. Levitov, see [56]. He proposed an “energetic” approach to the phyllotaxis, suggesting that the development of a plant is connected with an effective motion along the *geodesics* on the surface associated with the energetic relief of growing plant.

Intriguingly, the energetic relief considered in [56] coincides with the relief of the function $f(z)$ depicted in Fig. 8 (to the best of our knowledge, the Dedekind η function has never been discussed in this context). The black arcs drawn in Fig. 8a,b are parts of semicircles orthogonal to the boundaries and represent the geodesics of the surface $f(z)$, which are *open level lines* (by the symmetry, there are only two such level lines in Fig. 8a and three in Fig. 8b). The relief of the function $f(z)$ shown in the Fig. 8a emerged previously in course of

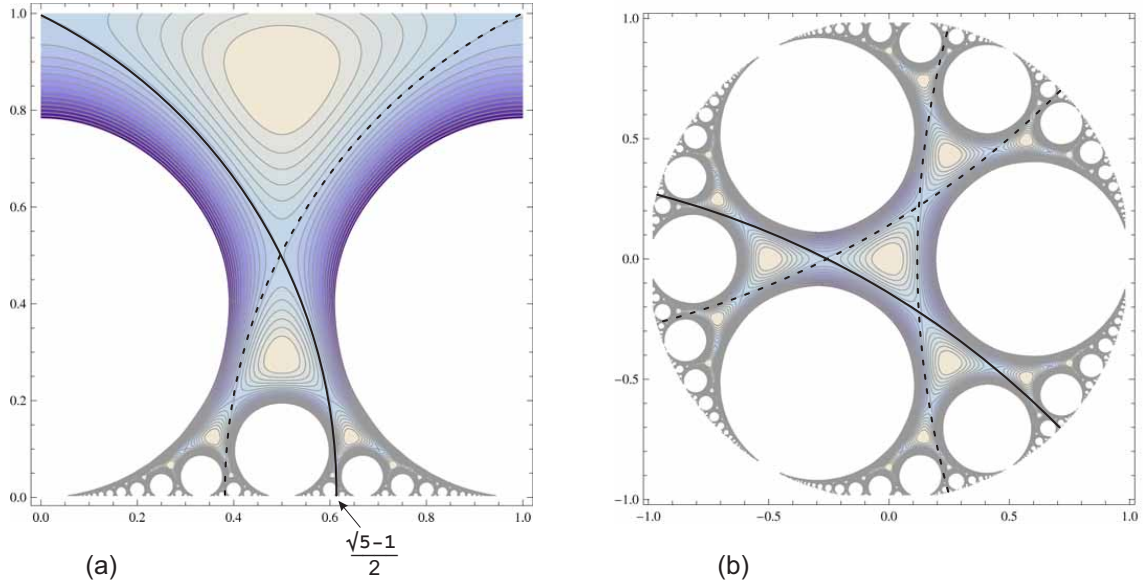


Figure 8: Cut ($f(z) \in [0.985, 1.0]$) of the contour plot of the function $f(z)$ (see (39)): a) the representation in the upper half-plane, $z = x + iy$; b) the representation in the unit disc obtained by the conformal mapping of the upper half-plane to the unit disc.

investigation of metric properties of ultrametric spaces [47]. The best rational approximation of the boundary point, $\frac{\sqrt{5}-1}{2}$, at which the solid geodesics in Fig. 8a terminates, is given by the continued fraction expansion

$$\frac{\sqrt{5}-1}{2} = \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \dots}}}, \quad (50)$$

cut at some level k . Cutting (50) at sequential k , we get the set of fractions constituting the interlacing Fibonacci sequences, $\{F_k\}$, in nominators and denominators:

$$\frac{1}{1}, \frac{1}{2}, \frac{2}{3}, \frac{3}{5}, \frac{5}{8}, \dots, \frac{F_{k-1}}{F_k}, \frac{F_k}{F_{k+1}}$$

Thus, the normalized η -function, $f(z)$, plays a role of the energy relief of a growing plant.

The Fibonacci numbers are the benchmarks, which fix the connection between ultrametric and hyperbolic geometry, since, on one hand they appear as discrete symmetries of the hyperbolic space, and on the other hand, being a subset of p -adic numbers, they uniquely parameterize the distances in the ultrametric space [57]. In the Appendix we summarize some facts which elucidate the relation between ultrametricity and hyperbolic geometry.

Acknowledgments

We are grateful to Y. I. Manin and Y. Fyodorov for valuable discussions. SN acknowledges the support of the IRSES DIONICOS grant, VA was supported by the Higher School

of Economics program for Basic Research, PLK thanks the Galileo Galilei Institute for Theoretical Physics and the INFN for support during the completion of this work.

ON ULTRAMETRICITY AND HYPERBOLIC GEOMETRY. MODULAR PARISI MATRIX

The embedding of a Cayley tree \mathcal{C} into the metric space is called "isometric" if \mathcal{C} covers that space, preserving all angles and distances. The Cayley tree \mathcal{C} isometrically covers the surface of constant negative curvature (the Lobachevsky plane) \mathcal{H} . One of possible representations of \mathcal{H} , known as a Poincaré model, is the upper half-plane $\text{Im } z > 0$ of the complex plane $z = x + iy$ endowed with the hyperbolic metric $ds^2 = \frac{dx^2 + dy^2}{y^2}$.

In [47] we have constructed the "continuous" analog of the standard 3-branching Cayley tree by means of modular functions and have analyzed the structure of the barriers separating the neighboring valleys. In particular, we have shown that these barriers are ultrametrically organized. The main ingredient of our construction is the function $f(z)$ defined in (39). The normalization constant C is chosen to fix the maximal value of $f(z)$ equal to 1: $0 < f(z) \leq 1$ for any z in the upper half-plane $\text{Im } z > 0$. The function $f(z)$ has the following property: all the solutions of the equation $f(z) - 1 = 0$ define all the coordinates of the 3-branching Cayley tree isometrically embedded into the upper half-plane $\mathcal{H}(z|\text{Im } z > 0)$ with hyperbolic metric. The density plot of the function $f(z)$ in the region $\{0 \leq \text{Re } z \leq 1, 0.04 \leq \text{Im } z \leq 1\}$ and its 3D relief are shown in the Fig. 9a,b. The Fig. 9b clearly displays the "continuous" tree-like structure of hills separated by the valleys. Consider now the function

$$v_y(x) = \sqrt{-\ln f(x + iy)}; \quad \text{Im } z > 0 \quad (51)$$

The typical shape of $v_{y=\text{const}}(x)$ is shown in the figure Fig. 9c for $y = 0.001$. The function $v_y(x)$ at any fixed y obeys the strong triangle inequality condition and hence possesses the ultrametric organization of barriers separating the valleys. At $y \rightarrow 0^+$ the function $v(x)$ exhibits the "linearity" in sequentially increasing and decreasing barriers similar to the well-known Thomae ("popcorn") function [58].

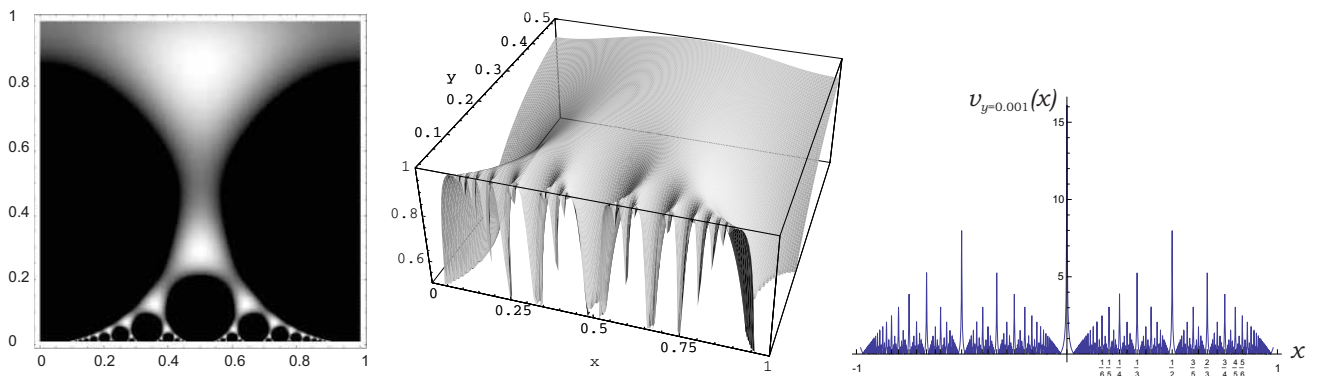


Figure 9: (a) Density plot of the function $f(z)$ in the rectangle $\{0 \leq \text{Re } z \leq 1, 0.01 \leq \text{Im } z \leq 1\}$; (b) 3D relief of the function $f(z)$ in the rectangle $\{0 \leq \text{Re } z \leq 1, 0.04 \leq \text{Im } z \leq 0.5\}$; (c) Plot of the function $v_y(x)$ at $y=0.001$.

The function $v_y(x)$, defined on the interval $0 < x < 1$, has the properties borrowed from the structure of the underlying modular group $PSL(2, \mathbb{Z})$ acting in the half-plane $\mathcal{H}(z|y > 0)$ [59, 60]. In particular: (i) the local maxima of $v_y(x)$ are located at the rational points; (ii) the highest barrier on a given interval $\Delta x = [x_1, x_2]$ is located at a rational point $\frac{p}{q}$ with the lowest denominator q . On a given interval $\Delta x = [x_1, x_2]$ there is only one such point. Locations of barriers with consecutive heights on the interval Δx are organized according to the group operation $\frac{p_1}{q_1} \oplus \frac{p_2}{q_2} = \frac{p_1+p_2}{q_1+q_2}$, constituting the Farey sequence discussed in the main text of the paper.

Let us mention finally that using the function $v_y(x)$ we can construct the "modular Parisi matrix", V_y , which for first 3 hierarchical levels is shown below:

$$V_y = \begin{array}{c} \begin{array}{|c|c|c|c|} \hline 0 & v_y\left(\frac{1}{4}\right) & v_y\left(\frac{1}{3}\right) & v_y\left(\frac{1}{3}\right) \\ \hline v_y\left(\frac{1}{4}\right) & 0 & v_y\left(\frac{1}{3}\right) & v_y\left(\frac{1}{3}\right) \\ \hline v_y\left(\frac{1}{3}\right) & v_y\left(\frac{1}{3}\right) & 0 & v_y\left(\frac{2}{5}\right) \\ \hline v_y\left(\frac{1}{3}\right) & v_y\left(\frac{1}{3}\right) & v_y\left(\frac{2}{5}\right) & 0 \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) \\ \hline v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) \\ \hline v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) \\ \hline v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) & v_y\left(\frac{1}{2}\right) \\ \hline \end{array} & \begin{array}{|c|c|c|c|} \hline 0 & v_y\left(\frac{3}{5}\right) & v_y\left(\frac{2}{3}\right) & v_y\left(\frac{2}{3}\right) \\ \hline v_y\left(\frac{3}{5}\right) & 0 & v_y\left(\frac{2}{3}\right) & v_y\left(\frac{2}{3}\right) \\ \hline v_y\left(\frac{2}{3}\right) & v_y\left(\frac{2}{3}\right) & 0 & v_y\left(\frac{3}{4}\right) \\ \hline v_y\left(\frac{2}{3}\right) & v_y\left(\frac{2}{3}\right) & v_y\left(\frac{3}{4}\right) & 0 \\ \hline \end{array} \\ \hline \end{array} \quad (52)$$

The matrix (52) can be considered as a "continuous" analog of the Parisi matrix and could play a role in problems where the ultrametricity should be consistent with metric structure of the target space.

-
- [1] K. Hensel, Jahresbericht der Deutschen Mathematiker-Vereinigung **6**, 83 (1897).
 - [2] N. Koblitz, *p-adic Numbers, p-adic Analysis, and Zeta-Functions* (New York: Springer, 1984).
 - [3] A. M. Robert, *A Course in p-adic Analysis* (New York: Springer, 2000).
 - [4] V. S. Vladimirov, I. V. Volovich and E. I. Zelenov, *p-Adic Analysis and Mathematical Physics* (New York: Word Scientific, 1995).
 - [5] V. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs, vol. 33 (Providence, RI: AMS, 1990); V. Berkovich, *Integration of One-forms on p-adic Analytic Spaces*, Annals of Mathematics Studies **162** (Princeton: Princeton University Press, 2006).
 - [6] P. G. O. Freund and M. Olson, Non-archimedean strings, Phys. Lett. B **199**, 186 (1987).
 - [7] P. G. O. Freund, and E. Witten, Adelic string amplitudes, Phys. Lett. B **199**, 191 (1987).

- [8] L. Brekke, P. G. O. Freund, M. Olson, and E. Witten, Non-archimedean string dynamics, Nucl. Phys. B **302**, 365 (1988).
- [9] B. Dragovich, A. Yu. Khrennikov, S. V. Kozyrev, and I. V. Volovich, On p -adic mathematical physics, p -Adic Numbers, Ultrametric Analysis, and Applications. **1**, 1 (2009).
- [10] D. Harlow, S. H. Shenker, D. Stanford, and L. Susskind, Tree-like structure of eternal inflation: A solvable model, Phys. Rev. D **85**, 063516 (2012).
- [11] V. Vanchurin, Continuum of discrete trajectories in eternal inflation, Phys. Rev. D **91**, 023511 (2015).
- [12] R. Rammal, G. Toulouse, and M. Virasoro, Ultrametricity for physicists, Rev. Mod. Phys. **58**, 765 (1986).
- [13] G. Parisi, A sequence of approximated solutions to the S-K model for spin glasses, J. Phys. A **13**, L115 (1980); G. Parisi, The order parameter for spin glasses: a function on the interval 0-1, J. Phys. A **13**, 1101 (1980).
- [14] M. Mézard, G. Parisi, and M. Virasoro, *Spin glass theory and beyond* (World Scientific, Singapore, 1987).
- [15] M. Talagrand, Replica Symmetry Breaking and Exponential Inequalities for the Sherrington-Kirkpatrick Model, Ann. Probab. **28**, 1018 (2000); F. Guerra and F.L. Toninelli, The Thermodynamic Limit in Mean Field Spin Glass Models, Commun. Math. Phys. **230**, 71 (2002).
- [16] D. Sherrington and S. Kirkpatrick, Solvable Model of a Spin-Glass, Phys. Rev. Lett. **35**, 179 (1975).
- [17] H. Frauenfelder, The Connection Between Low-Temperature Kinetics and Life, in *Protein Structure: Molecular and Electronic Reactivity*, Ed. by R. Austin et al. (New York: Springer, 1987), pp. 245.
- [18] H. Frauenfelder, Complexity in Proteins, Nat. Struct. Biol. **2**, 821 (1995).
- [19] V. A. Avetisov, A. Kh. Bikulov, and A. P. Zubarev, Ultrametric Random Walk and Dynamics of Protein Molecules, Proc. Steklov Inst. Math. **285**, 3 (2014).
- [20] V.A. Avetisov, L. Nazarov, S. K. Nechaev, and M. V. Tamm, A statistical model of intra-chromosome contact maps, Soft Matter **11**, 1019 (2015)
- [21] A. P. Zubarev, On stochastic generation of ultrametrics in high-dimension Euclidean spaces, p -Adic Numbers, Ultrametric Anal. Appl. **6**, 150 (2014).
- [22] G. J. Rodgers and A. J. Bray, Density of states of a sparse random matrix, Phys. Rev. B **37**, 3557 (1988).
- [23] T. Rogers, I. P. Castillo, R. Kühn, and K. Takeda, Cavity approach to the spectral density of sparse symmetric random matrices, Phys. Rev. E **78**, 031116 (2008).
- [24] Y. V. Fyodorov and A. D. Mirlin, On the density of states of sparse random matrices, J. Phys. A: Math. Gen. **24**, 2219 (1991).
- [25] S. N. Evangelou and E. N. Economou, Spectral density singularities, level statistics, and localization in a sparse random matrix ensemble, Phys. Rev. Lett. **68**, 361 (1992).
- [26] M. Bauer and O. Golinelli, Random Incidence Matrices: Moments of the Spectral Density, J. Stat. Phys. **103** 301 (2001).
- [27] G. Semerjian and L. F. Cugliandolo, Sparse random matrices: the eigenvalue spectrum revisited, J. Phys. A: Math. Gen. **35** 4837 (2002).
- [28] R. Kühn, Spectra of sparse random matrices, J. Phys. A: Math. Theor. **41**, 295002 (2008).
- [29] O. Rojo and R. Soto, The spectra of the adjacency matrix and Laplacian matrix for some balanced trees, Linear Alg. Appl. **403**, 97 (2005).
- [30] O. Rojo and M. Robbiano, On the spectra of some weighted rooted trees and applications,

- Linear Alg. Appl. **420**, 310 (2007).
- [31] O. Rojo and M. Robbiano, An explicit formula for eigenvalues of Bethe trees and upper bounds on the largest eigenvalue of any tree, Linear Alg. Appl. **427**, 138 (2007).
 - [32] S. Bhamidi, S. N. Evans, and A. Sen, Spectra of large random trees, J. Theor. Probab. **25**, 613 (2012).
 - [33] S. Janson, T. Łuczak, and A. Rucinski, *Random Graphs* (New York: John Wiley & Sons, 2000).
 - [34] P. L. Krapivsky, S. Redner and E. Ben-Naim, *A Kinetic View of Statistical Physics* (Cambridge: Cambridge University Press, 2010).
 - [35] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics: A Foundation for Computer Science* (Reading, Mass.: Addison-Wesley, 1989).
 - [36] P. Flajolet and R. Sedgewick, *Analytic Combinatorics* (Cambridge University Press, New York, 2009).
 - [37] E. Ben-Naim and P. L. Krapivsky, Unicyclic components in random graphs, J. Phys. A **37**, L189 (2004); E. Ben-Naim and P. L. Krapivsky, Kinetic theory of random graphs: From paths to cycles, Phys. Rev. E **71**, 026129 (2005).
 - [38] M. Mobilia, P. L. Krapivsky, and S. Redner, Kinetic anomalies in addition-aggregation processes, J. Phys. A **36**, 4533 (2003).
 - [39] E. Ben-Naim and P. L. Krapivsky, Kinetics of ring formation, Phys. Rev. E **83**, 061102 (2011).
 - [40] E. Ben-Naim and P. L. Krapivsky, Dynamics of random graphs with bounded degrees, J. Stat. Mech. P11008 (2011).
 - [41] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers* (Oxford: Oxford University Press, 1979).
 - [42] L. R. Ford, Fractions, Amer. Math. Monthly **45**, 586 (1938); H. S. M. Coxeter, The problem of Apollonius, Amer. Math. Monthly **75**, 5 (1968).
 - [43] Y. Georgelin, T. Masson, and J.-C. Wallet, Modular groups, visibility diagram and quantum Hall effect, J. Phys. A: Math. Gen. **30**, 5065 (1997).
 - [44] I. M. Lifshitz, S. A. Gredeskul, and L. A. Pastur, *Introduction to the theory of disordered systems* (Wiley-Interscience, 1988).
 - [45] W. Kirsch and I. Veselic, Lifshitz Tails for a Class of Schrödinger Operators with Random Breather-Type Potential, Lett. Math. Phys. **94**, 27 (2010).
 - [46] K. Chandrasekharan, *Elliptic Functions* (Berlin: Springer, 1985).
 - [47] S. Nechaev and O. Vasilyev, Metric structure of ultrametric spaces, J. Phys. A: Math. Gen. **37**, 3783 (2004).
 - [48] B. Bollobás, *Modern Graph Theory*, (Berlin: Springer, 1998, corrected ed. 2013)
 - [49] V. Trifonov, L. Pasqualucchi, R. Dalla-Favera, and R. Rabadan, Fractal-like distributions over the rational numbers in high-throughput biological and clinical data, Scientific Reports, **1**, 191 (2011)
 - [50] D. Zagier, Vassiliev invariants and a strange identity related to the Dedekind eta-function, Topology **40**, 945 (2001).
 - [51] R. Lawrence and D. Zagier, Modular forms and quantum invariants of 3-manifolds. Asian J. of Math. **3**, 93 (1999).
 - [52] D. Zagier, Quantum modular forms, Clay Mathematics Proceedings **12**, 1 (2010).
 - [53] S. Nechaev and R. Voituriez, On the plant leaves boundary, “jupe à godets” and conformal embeddings, J. Phys. A: Math. Gen. **34**, 11069 (2001).
 - [54] B. Audoly and Y. Pomeau, *Elasticity and Geometry: From hair curls to the non-linear re-*

- sponse of shells* (Oxford: Oxford Univ. Press, 2010).
- [55] I. Adler, *Solving the Riddle of Phyllotaxis: Why the Fibonacci Numbers and the Golden Ratio Occur On Plants* (Singapore: World Scientific, 2012).
 - [56] L. S. Levitov, Energetic Approach to Phyllotaxis, Europhys. Lett. **14**, 533 (1991).
 - [57] V. A. Avetisov and S. K. Nechaev, Chaotic Hamiltonian systems: Survival probability, Phys. Rev. E **81**, 046211 (2010).
 - [58] K. Beanland, J.W. Roberts, and C. Stevenson, *Modifications of Thomae's Function and Differentiability* Amer. Math. Monthly **116**, 531 (2009).
 - [59] W.Magnus, *Noneuclidean Tessellations and their Groups* (Acad.Press: London, 1974)
 - [60] A.F. Beardon, *The Geometry of Discrete Groups* (Springer: Berlin, 1983)