Notes on Integrable Probability

Galobel Wang
The University of Hong Kong

May 2025

Abstract

This article consists of some notes taken by the author while studying integrable probability. References: [1], [2], [3], [4]

Contents

1 Introduction

Integrable probability is a fascinating subfield of probability theory that centers on probabilistic models endowed with special mathematical structures, allowing for exact solutions through precise mathematical techniques. These models often connect to the notion of integrable systems, characterized by high symmetry or hidden mathematical structures. The term **integrable** is adapted from statistical physics, where it refers

to **exact solvability** or **integrable systems**, but in integrable probability, its meaning is broader and more flexible, encompassing a wider range of models. A defining characteristic of integrable probability is the ability to express transition probabilities or probability distributions using sophisticated mathematical tools such as **symmetric functions**, **orthogonal polynomials**, or **contour integrals of special functions**. This property not only sets integrable models apart but also enables rigorous **asymptotic analysis**, making it possible to derive their limiting distributions with precision.

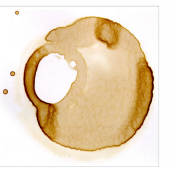
The scope of integrable probability spans a diverse array of models, including both discrete and continuous types:

- **Discrete models**: Examples include the geometric Directed Last Passage Percolation (DLPP), discrete Totally Asymmetric Simple Exclusion Process (TASEP) or Asymmetric Simple Exclusion Process (ASEP), and the stochastic six-vertex model.
- **Continuous models**: These encompass the Kardar-Parisi-Zhang (KPZ) equation, random polymer models, and continuous TASEP/ASEP.

Despite their differences, these models share a common thread: their transition probabilities or distributions can be articulated through the aforementioned mathematical expressions, yielding insights into their limiting behaviors.

The concept of **integrability** in this context is notably flexible. For instance, within the **KPZ universality class**, some models that lack strict solvability have still been proven to converge to KPZ limits, broadening the traditional boundaries of integrable systems. The ultimate ambition of integrable probability is to construct a unified framework—akin to the **Central Limit Theorem**(CLT) — that comprehensively explains the limiting behaviors of **integrable**, **approximately integrable**, and even **non-integrable models**. This vision underscores the field's potential to bridge diverse probabilistic phenomena under a single theoretical umbrella.

1.1 Universality in Random Systems



In a coffee stain, before it is dry, particle diffuse and eventually stick to the boundary. This is why the edges of the stain are darker.

The concept of **universality** arose in the context of statistical mechanics as a consequence of the study of critical phenomena. To understand what a critical phenomenon

is, we consider a bar magnet. We describe its capacity of picking up thumbtacks with an order parameter M. As we heat the system, M decreases and eventually, at a certain critical temperature T_c, it reaches zero, which means the bar loses its magnetization. The special value of the parameter at which the system changes its phase is the critical point. For systems that exhibit universality, the closer the parameter is to its critical value, the less sensitively the order parameter depends on the details of the system. The order parameter is well approximated by

$$M \sim |T - T_c|^{\alpha},$$

where the exponent α is the critical exponent. It was found empirically that one can partition critical systems into **universality classes**. In the last decade the concept of universality class has been investigated and exploited and now it plays a central role in probability and mathematical physics.

1.1.1 Gaussian universality class

The number of heads H in N fair coin flips is given exactly by the Binomial distribution:

$$\mathrm{P}(H = n) = 2^{-N}\binom{N}{n}. \tag{1.1}$$

- Law of large numbers [Bernoulli 1713]:

$$\lim_{N\to\infty} \frac{H}{N} = \frac{1}{2}. \tag{1.2}$$

- Central limit theorem [de Moivre 1733, Laplace 1812]:

$$\lim_{N\to\infty} \mathrm{P}\left(H < \frac{1}{2}N + \frac{1}{2}\sqrt{Nx}\right) = \int_{-\infty}^{x} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy. \tag{1.3}$$

The proof of the central limit theorem shows some basic ideas in universality classes.

- Asymptotics of n! [de Moivre 1721, Stirling 1729]:

$$n! = \Gamma(n+1) = \int_0^{\infty} e^{-t} t^n dt = n^{n+1} \int_0^{\infty} e^{nf(z)} dz \tag{1.4}$$

  where $f(z) = \log z - z$ and the last equality is from the change of variables $t = nz$.

- The integral is dominated, as n grows, by the maximal value of $f(z)$ on the interval $[0, \infty]$. This occurs at $z = 1$, thus expanding $f(z) \approx -1 - \frac{(z-1)^2}{2}$, and plugging

this into the integral yields the final expansion:

$$n! \approx n^{n+1} e^{-n} \sqrt{\frac{2\pi}{n}} \tag{1.5}$$

This general route: (1.1) writing exact formulas for probabilities in terms of integrals; (1.2) performing asymptotics analysis.

The universality of the Gaussian distribution was not demonstrated until [Lyapunov 1901].

The central limit theorem (CLT) showed that the exact nature of coin flipping is immaterial—any sum of independent identically distributed (iid) random variables with finite mean and variance will demonstrate the same limiting behavior.

Theorem 1.1 (central limit theorem)**.** *Let $X_1, X_2, \ldots$ be i.i.d random variables of finite mean m and variance ν.Then for all $x \in \mathbb{R}$,*

$$\lim_{N\to\infty} \mathbb{P}\left(\sum_{i=1}^{N} X_i < mN + \nu\sqrt{N}x\right) = \int_{-\infty}^{x} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy \tag{1.6}$$

The above theorem has an equivalent statement.

Theorem 1.2 (central limit theorem)**.** *Let $X_1, X_2, \ldots$ be a sequence of independent, identically distributed random variables with mean μ and variance $\sigma^2 < \infty$. Let $S_n := \sum_{k=1}^{n} X_k$. Then*

$$\lim_{n\to\infty} P\left(\frac{S_n - n\mu}{\sigma\sqrt{n}} \leq s\right) = \int_{-\infty}^{s} \frac{e^{-x^2/2}}{\sqrt{2\pi}}\, dx. \tag{1.1.2}$$

This means that, despite the model dependent features as the mean and the variance, any sum of i.i.d random variables with finite variance will show the same limiting behavior, described by the normal distribution, and fluctuations around the mean value of order $n^{1/2}$. Physical and mathematical systems accurately described in terms of Gaussian statistics are said to be in the Gaussian universality class.

Theorem 1.2 illustrates well the universal and the non-universal quantities. The model-dependent and thus non-universal quantities are μ and σ^2, while the $1/2$ exponent in the normalization as well as the normal distribution are universal.

1.1.2 Random versus ballistic deposition models

In this part we will be dealing with models related to stochastically growing interfaces. Let us starts with two simple models, which however shows very different limiting behaviors.

In the **random deposition model** unit blocks fall independently and in parallel on $\mathbb{Z}$ after an exponentially distributed waiting time of parameter 1 (see Figure 1). Due to the memoryless property of the exponential distribution, this model is a Markov

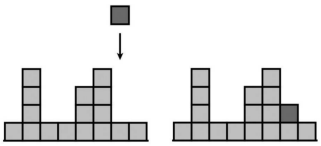


Figure 1: random deposition model (a)

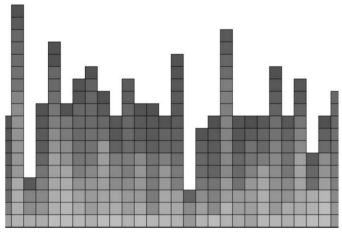


Figure 2: random deposition model (b)

process; its evolution depends only on the present and not on the history. Each column evolves independently and is a Poisson process.

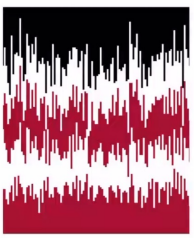
- Exponential distribution of rate λ (mean $\frac{1}{\lambda}$): $\mathbb{P}(X > x) = e^{-\lambda x}$.
- Memoryless (Markov), so growth depends only on the present state.
- Gaussian behavior since each column is a sum of i.i.d random variables.

Denote by $h(x,t)$ the height function that gives the value of the height of the column on the site x at time t. We assume that $h(x,0) = 0$ for all x. Random waiting times $\omega_{x,i}$: the time for the i-th block in column x to fall. For any n, the event $h(t,x) < n$ is equivalent to $\sum_{i=1}^{n} \omega_{x,i} > t$. Since the $\omega_{x,i}$ are i.i.d, the law of large numbers and central limit theory apply here. Assume $\lambda = 1$,

$$\lim_{t\to\infty} \frac{h(x,t)}{t} = 1,$$

and

$$\frac{h(x,t) - t}{\sqrt{t}} \xrightarrow{d} N(x).$$



According to the Law of Large Numbers and the Central Limit Theorem, we know that the random deposition model has the following properties:

- Linear growth (known speed)
- $t^{\frac{1}{2}}$ fluctuations with Gaussian limit (CLT)
- No spatial correlation

In general, **fluctuations and transversal correlation exponents**, as well as **limiting distributions** constitutes the description of a universality class and all models which match these limiting behaviors are said to lie in the same universality class.

1.1.3 Edwards-Wilkinson universality class

Compared to the random deposition model, it can be observed from the image that the growth interface of the Edwards-Wilkinson universality class is smoother and more stable.

- Relaxation: when a block falls over site x, it lands on either x or any of its two nearest neighbors, whichever has the lowest height (choosing, say, uniformly in case of ties).
- Linear growth.
- $t^{\frac{1}{4}}$ fluctuations with Gaussian limit (CLT).
- Spatial correlation of order $t^{\frac{1}{2}}$.
- (time : space : fluctuation) is (4 : 2 : 1).

In other words, if a probabilistic model satisfies the condition **(time : space : fluctuation) = (4 : 2 : 1)**, it belongs to the **Edwards-Wilkinson universality class**.

1.1.4 Ballistic deposition (or sticky block) model

It is enough to slightly modify the rules of the growth of the interface to lose the Gaussian behavior. Consider the same model, but now, instead of falling on the ground

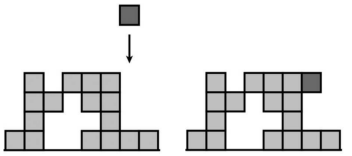


Figure 3: ballistic deposition model (a)

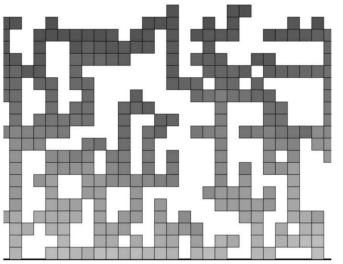


Figure 4: ballistic deposition model (b)

or the interface, a new block sticks to the first edge (see Figure 3): this is known as the ballistic deposition model, introduced by Vold in 1959. This change in the evolution rules turns into large time effects: the interface grows faster than in the random deposition model (the value of the velocity is still unknown) and simulations (see Figure 4) show that the height function has smaller fluctuations, on the scale $t^{\frac{1}{3}}$, and demonstrate nontrivial correlations on the transversal scale of $t^{\frac{2}{3}}$. Moreover the rescaled height does not converge anymore to the Gaussian distribution.

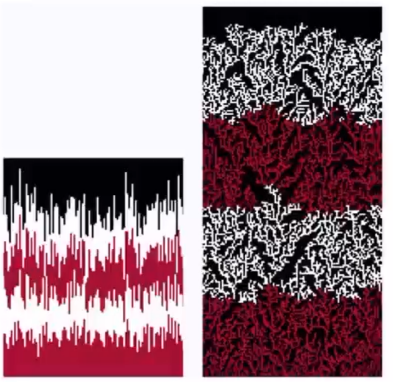


Figure 5: comparison

Compared to the random deposition model, it can be observed from the image that the growth interface of the ballistic deposition model is smoother and more stable (see Figure 5), Different colors represent different time intervals.

Ballistic deposition model have the following properties:

- Generally linear growth (unknown speed): t^1
- Conjectural $t^{\frac{2}{3}}$ spatial correlation
- Conjectural $t^{\frac{1}{3}}$ fluctuations
- $\lim_{t\to\infty} t^{-\frac{1}{3}} \left(h(c_1 t, c_2 t^{\frac{2}{3}} x) - c_3 t \right) \stackrel{d}{=} \mathcal{A}(x)$ (CLT, where $\mathcal{A}$ corresponding to Gaussian distribution in Gaussian universality and $\mathcal{A}$ only depends on the initial condition)
- 1:2:3 KPZ scaling (fluctuation : space : time)

1.1.5 KPZ universality class

By extracting the characteristics of the Ballistic deposition model (because ballistic deposition model is not integrable), find some similar alternative models, then prove that those alternative models satisfy KPZ scaling, and finally use the special models to determine the distribution $\mathcal{A}$.

The ballistic deposition model is believed to belong to the **Kardar-Parisi-Zhang (KPZ)** universality class. A model is in this universality class if the following properties are satisfied:

- Locality: Height function change depends only on neighboring heights (the evolution of the height function depends only the local (in space and time) configuration of the interface).
- Smoothing: Large valleys are quickly filled (there is a smoothing mechanism implying that deep holes are rapidly filled to smooth the interface such that there exists a non-random macroscopic limit shape)

$$h_{ma}(\xi, \tau) = \lim_{t\to\infty} \frac{h(\xi, t, \tau)}{t}.$$

- Non-linearity of speed of growth: vertical effective growth rate depends non-linearly on local slope (Macroscopically the growth of the interface will evolve deterministically according to some PDE. Since the growth rules are local and locally one will see a given slope only, one will get $\partial_\tau h_{ma} = v(u = \nabla h_{ma})$. In order to be in the KPZ class, we need to have the condition $v''(u) \neq 0$. This property implies irreversibility of the model).
- Space-time independent noise: Growth is drive by noise which quickly decorrelates in space/time and is not heavy tailed.

The KPZ universality class was introduced in the context of studying the motion of growing interfaces in 1986 in a paper of Kardar, Parisi, and Zhang, where they studied

a continuum stochastically growing height function $h(x,t)$ given in terms of a stochastic PDE, known as the KPZ equation:

$$\frac{\partial h}{\partial t}(x,t) = \nu \frac{\partial^2 h}{\partial x^2}(x,t) + \lambda \left(\frac{\partial h}{\partial x}(x,t) \right)^2 + \sqrt{D}\xi(x,t),$$

where $\xi(x,t)$ is Gaussian space-time white noise, $\lambda, \nu \in \mathbb{R}, D > 0$ (the stationary solution is formally a two-sided Brownian motion). This equation is ill-posed, since the nonlinearity does not make sense in this case. The equation contains of course the key features: local random growth (the noise ξ), the smoothing (from the Laplacian), the non-linearity (from the square of the gradient).

1.1.6 Corner growth model

So far, the ballistic deposition model could not be investigated analytically, but there is a class of systems that is integrable. For an integrable probabilistic system, it is possible to compute concise formulas for averages of a class of observables; furthermore, taking limits of the system, observables and formulas, it is possible to access detailed descriptions of universal classes. We will focus on a few examples in the Kardar-Parisi-Zhang class.

Consider an interface modeled by a height function $h(x,t)$, $x, t \in \mathbb{Z}$, with $h(x \pm 1, t) - h(x,t) \in \{-1, 1\}$. WLOG set $h(0,0) = 0$. The height function evolves according to the following dynamics: each local minimum turns into a local maximum after an exponentially distributed waiting time of parameter 1 (see Figure 6). Particularly interesting are the cases of two initial configurations: wedge initial condition, which means that $h(x,0) = |x|$, and flat initial condition, which means that $h(x,0)$ is a sawtooth function between 0 and 1.

- Height function $h(t,x)$: continuous, piecewise linear, and composed of $\sqrt{2}$-length line increments of slope +1 or -1, changing value at integer x.
- The height function evolves according to the Markovian dynamics that each local minimum of h (looking like $\vee$) turns into a local maximum (looking like $\wedge$) according to an exponentially distributed waiting time. This happens independently for each minimum.
- This change in height function can also be thought of as adding boxes (rotated by 45°).

There is also an equivalent formulation of this growth model. Project the interface to a straight line and put "particles" at projections of unit segments of slope -1 and "holes" at projections of segments of slope +1 (see Figure 7 left). Now each particle independently jumps to the right after an exponential waiting time (put it otherwise,

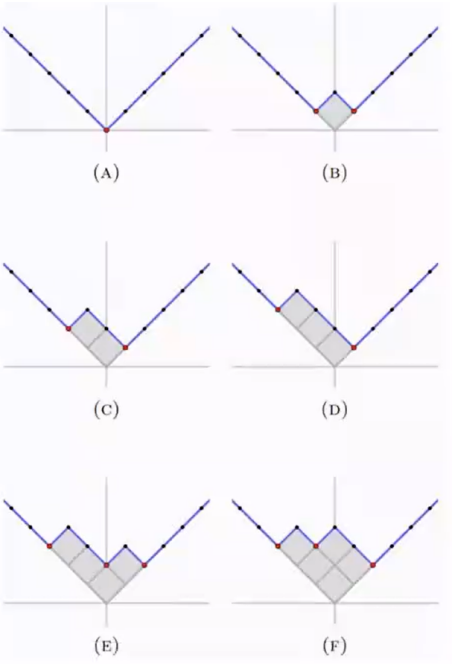


Figure 6: Evolution of the height function

each particle jumps with probability dt in each very small time interval $[t, t+dt]$) except for the exclusion constraint: Jumps to the already occupied spots are prohibited. This is a simplified model of a one-lane highway which is known under the name of Totally Asymmetric Simple Exclusion Process (TASEP), cf.

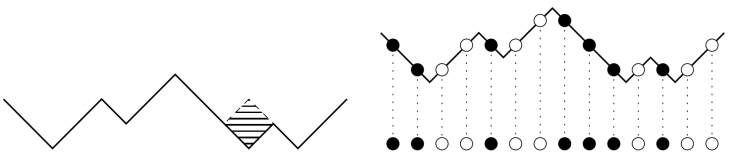


Figure 7: Broken line with slopes ± 1, local minimum where a box can be added, and correspondence with particle configurations on $\mathbb{Z}$

The macroscopic limit shape for wedge initial condition is a parabola continued by two straight lines:

$$h_{ma}(\xi) = \begin{cases} \frac{1}{2}(1+\xi^2), & \text{for } |\xi| \le 1 \\ |\xi|, & \text{for } |\xi| \ge 1. \end{cases}$$

Johansson proved large time results for the rescaled height function.

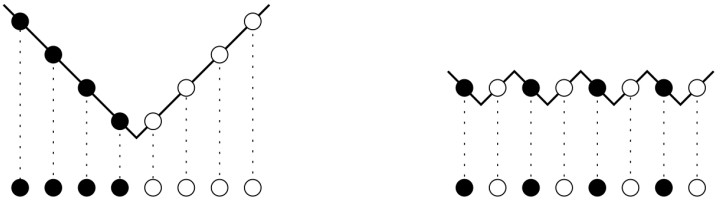
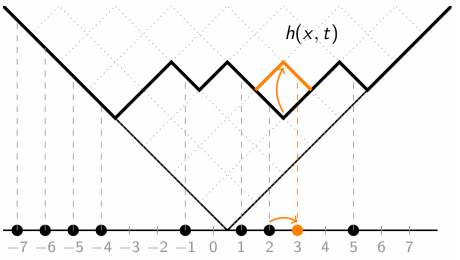


Figure 8: Wedge and flat initial conditions: broken lines and corresponding particle configurations.



Theorem 1.3 (Johansson)**.** *For wedge initial condition, for any* $\xi \in (-1,1)$*,*

$$\lim_{t\to\infty} \mathbb{P}\left(\frac{h(\xi t,t)-c_1(\xi)t}{-c_2(\xi)t^{1/3}} \le s\right) = F_2(s),$$

with c_1, c_2 *model-dependent constants.*

The function F_2 is known as the GUE Tracy-Widom distribution, first discovered in random matrices. For flat initial condition the macroscopic shape is simply $h_{ma}(\xi) = \frac{1}{2}$. An analogous result for the limit distribution stands also for this case.

Theorem 1.4 (Borodin-Ferrari-Prähofer-Sasamoto)**.** *For flat initial condition, for any* $\xi \in (-1,1)$*,*

$$\lim_{t\to\infty} \mathbb{P}\left(\frac{h(\xi t,t)-c_3(\xi)t}{-c_4(\xi)t^{1/3}} \le s\right) = F_1(2s),$$

with c_3, c_4 *model-dependent constants.*

Also the function F_1 was first observed in the random matrix context and it is known as GOE Tracy-Widom distribution. An interesting fact to observe is that, although the scaling exponent are invariant, the limit distribution depends on the initial condition.

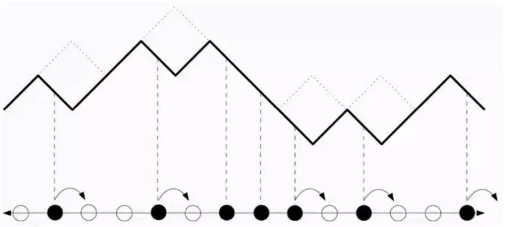
Here $F_1(s)$ and $F_2(s)$ are distributions from random matrix theory, known under the name of Tracy-Widom distributions. They are the limiting distributions for the largest eigenvalues in Gaussian Orthogonal Ensemble and Gaussian Unitary Ensemble of random matrices (which are the probability measures with density proportional to $\exp(-\mathrm{Trace}(X^2))$ on real symmetric and Hermitian matrices, respectively).

These two theorems give the conjectural answer for the whole "universality class"

of $2d$ random growth models, which is usually referred to as the KPZ (Kardar-Parisi-Zhang) universality class. Comparing to the answer in the $1d$ case we see that the asymptotic behavior becomes more delicate — while scaling by $t^{\frac{1}{3}}$ is always the same, the resulting distribution may also depend on the "subclass" of our model. Also, conjecturally, the only two generic subclasses are the ones we have seen. They are distinguished by whether the global surface profile is locally curved or flat near the observation location.

1.1.7 Interacting particle systems

Totally Asymmetric Simple Exclusion Process (TASEP) model:



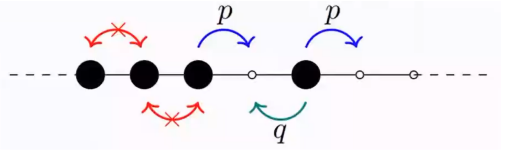
- KPZ fixed point (goal): temporal evolution of the interface

$$\lim_{\epsilon \to 0} \epsilon^{\frac{1}{2}} \left(h\left(c_1 \epsilon^{-\frac{3}{2}} t, c_2 \epsilon^{-1} x \right) - c_3 \epsilon^{-\frac{3}{2}} t \right) = \hat{h}(t, x).$$

- Obstacle 1: we only know how to compute the limit for two special choices of initial data.
- Obstacle 2: we can only do it for fixed time t.

Asymmetric simple exclusion process (ASEP) model: more generally, perturb the TASEP or study to what extent the TASEP model remains integrable.

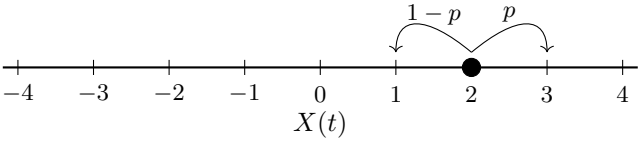
p and q represent the probabilities of particles moving to the right and to the left, respectively.



- $p > q$: KPZ universality class.
- $p = q$: Edwards-Wilkinson universality class.
- One-point distribution: GUE Tracy-Widom distribution.

- Multi-point and multi-time distribution: KPZ fixed point.
- $p - q \to 0$: KPZ equation.

Example 1.1. *Consider the simplest particle system.*



- *We can associate a difference operator*

$$\Delta f(x) = p\, f(x-1) + (1-p) f(x+1),$$

so that

$$\psi_t(x) := \mathbb{P}(X(t) = x)$$

solves

$$\psi_{t+1}(x) = \Delta \psi_t(x).$$

- *For any $z \in \mathbb{C} \setminus \{0\}$, functions $x \mapsto z^x$ are eigenfunctions. We may consider the Fourier transform*

$$\hat{f}(z) = \sum_{x \in \mathbb{Z}} f(x) z^x,$$

which can be inverted via

$$f(x) = \frac{1}{2\pi i} \oint_{|z|=1} \hat{f}(z) \frac{dz}{z^{1+x}}.$$

(isometry between $\ell^2(\mathbb{Z})$ and $L^2(\mathbb{T}, \frac{dz}{z})$)

Recall $\psi_0(x) = \mathbb{P}(X(0) = x)$, and consider the identity

$$\psi_0(x) = \frac{1}{2i\pi} \oint_{|z|=1} \widehat{\psi_0}(z) \frac{dz}{z^{1+x}}.$$

Acting t times with Δ on both sides, we obtain (recall $\psi_{t+1}(x) = \Delta \psi_t(x)$)

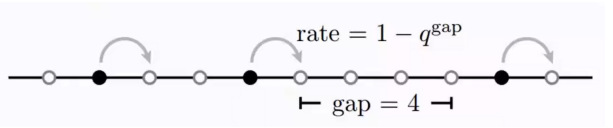
$$\mathbb{P}(X(t) = x) = \frac{1}{2i\pi} \oint_{|z|=1} \underbrace{\left(pz + (1-p)z^{-1}\right)^t}_{eigenvalue} \underbrace{\left(\sum_{y \in \mathbb{Z}} z^y \mathbb{P}(X(0) = y) \right)}_{\widehat{\psi_0}(z)} \frac{dz}{z^{1+x}}.$$

The last expression can be analyzed asymptotically using standard techniques of asymptotic analysis for contour integrals.

- *For more complicated systems, eigenfunctions will not be as simple as z^x but typically functions of many variables $z_1, z_2, z_3, \ldots$*
- *If we have now several particles jumping on $\mathbb{Z}$ with at most one particle per site and $p = 1$, all of this goes through, but the functions $x \mapsto z^x$ are replaced by so-called Grothendieck polynomials*

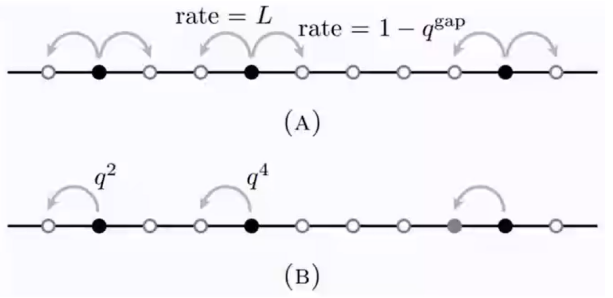
$$G_{\vec{x}}(z_1, \ldots, z_n) = \frac{\det\left(z_i^{x_j}(1 - 1/z_i)^{1-j}\right)_{i,j=1}^{n}}{\det\left(z_i^{n-j}\right)_{i,j=1}^{n}}.$$

q-TASEP: Consider another form of perturbation involving the deceleration mechanism.



- Cars slow down as they approach the one in front.
- Particles jump right according to independent exponential waiting times of rate $1 - q^{gap}$.
- KPZ class fluctuation behavior: GUE Tracy-Widom distribution.
- Method 1: Macdonald processes.
- Method 2: Bethe Ansatz.

q-push ASEP: Introduce a braking mechanism.



- The cascade effect of braking.
- Particles still jump right according to q-TASEP rules; however, now particles may also jump left after exponential rate L waiting times.

- When such a jump occurs, it prompts the next particle to the left to likewise jump left, with a probability given by q^{gap} where gap is the number of empty spaces between the original particle and its left neighbor.

1.1.8 Paths in random environment

Last passage percolation model: In the Corner Growth Model, we use the height function to characterize the process. Now we will use waiting times to describe this growth model.

Consider the model of growing interface just introduced starting from wedge initial configuration. There is an alternative way to describe the evolution of the height function, by taking the time at which a box appears on a local minimum, and call this time $T_{i,j}$ for $i, j \in \mathbb{N}$. Let indicate with $w_{i,j}$ the "waiting time" for the local valley at (i, j) to become a local maximum. Since a box can appear only once the blocks $(i-1, j)$ and $(i, j-1)$ have appeared, $T_{i,j}$ must satisfy the recursive relation $T_{i,j} = T_{i-1,j} \vee T_{i,j-1} + w_{i,j}$, and iterating

$$T_{i,j} = \max_{\pi:(1,1)\to(i,j)} \sum_{(k,l)\in\pi} w_{k,l}$$

where π are all steps made of consecutive steps of $(1, 0)$ or $(0, 1)$. This model is called last passage percolation with exponential weight. Another solvable model is last passage percolation with geometric weights, namely take $\mathbb{P}(w_{i,j} = k) = (1-q)q^k$ for $k \geq 0$. The exponential case can be recovered by the limit $q \to 1$ and setting

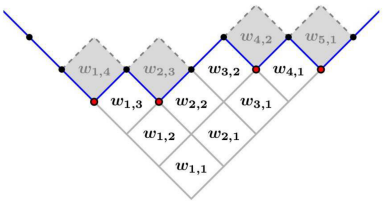


Figure 9: An example of last passage percolation

$k = t/(1-q)$. Another limit is obtained by letting $q \to 0$ and choosing the unit of i, j to be $\sqrt{2}$ instead of 1. If we draw one point each time that there is a $w_{i,j} = 1$ (higher values will not occur in the limit), then we are left with a Poisson point process with density 1 on $\mathbb{R}^2$. The paths can then be taken to taken to be directed paths which maximizes the number of Poisson points visited. Once can reformulate the problem as the one of finding the longest increasing subsequence of a random permutation, which was solved by Baik, Deift and Johansson:

Theorem 1.5. *Let σ be a uniformly distributed permutation of $\{1, \ldots, n\}$ and let $\ell_n(\sigma)$*

be the length of the longest increasing subsequence. Then

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{\ell_n(\sigma) - 2\sqrt{n}}{n^{1/6}} \le s\right) = F_2(s).$$

Here the macroscopic parameter is $\sqrt{n}$ and the fluctuations are $n^{1/3} = n^{1/6}$. In the next chapter we will consider this model with $\mathbb{P}(w_{i,j} = k) = (1 - a_i b_j)(a_i b_j)^k$, $k \ge 0$, for arbitrary family of parameters $\{a_i, b_j\}$ satisfying $0 < a_i, b_j < 1$. The reason is that having the parameters, the mathematical structure in the background is much more visible.

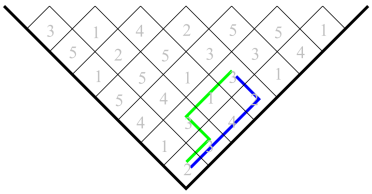
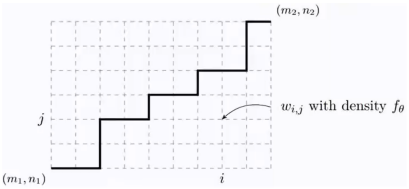


Figure 10: The quadrant filled with waiting times and two (out of $\binom{4}{1} = 4$ possibilities) directed paths joining $(1, 1)$ and $(4, 1)$.

Directed polymers in random environment



- Inverse-gamma distribution: $f(x) = \frac{1}{\Gamma(\theta)} x^{-\theta-1} e^{-\frac{1}{x}}$.
- π: up/right directed lattice path from point (m_1, n_1) to (m_2, n_2).
- Partition function:

$$Z(m_1, n_1; m_2, n_2) = \sum_{\pi:(m_1,n_1)\to(m_2,n_2)} \prod_{\pi} \omega_{i,j}.$$

- Free energy log Z demonstrated KPZ class fluctuations.
- Open problem: KPZ class fluctuation persists when the distribution of the $\omega_{i,j}$ is arbitrary.

OPEN problem: KPZ universality class fluctuation persists when the distribution of the $w_{x,y}$ is arbitrary?

Universality considerations make one believe that the limit behavior of the Last Passage Percolation time should not depend on the distribution of $w_{i,j}$ (if this distribution is not too wild), but we are very far from proving this at the moment. However, again, the Last Passage Percolation time asymptotics has been computed for certain distributions, e.g. for the exponential distribution.

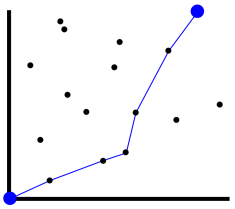


Figure 11: The Poisson point process in the first quadrant and a North-East path joining $(0,0)$ and (θ,θ) and collecting maximal number of points, which is 5 here.

The Hammersley Process:

- Degenerate LPP problem.
- Density 1 Poisson point process.
- **Hammersley problem**: what is the maximum number of Poisson points that can be collected by going from $(0,0)$ to (N,N) via an up-right path?

Compact set $A \subset \mathbb{R}^2_{\geq 0}$, let N_A be the number of particles falling in the set A.

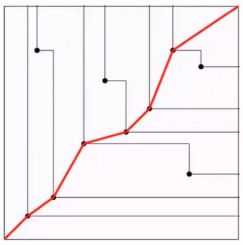
- N_A has the Poisson distribution with parameter $|A|$, the area of $|A|$. That is $P(N_A = n) = e^{-|A|}\frac{|A|^n}{n!}$, $n = 0, 1, 2, \ldots$.
- If $A_1, \ldots, A_k$ are pairwise disjoint compact sets, then the corresponding random variables $N_{A_1}, \ldots, N_{A_k}$ are independent.

Let us present another example, where the (conjecturally, universal) result can be rigorously proven. Consider the homogeneous, density 1 Poisson point process in the first quadrant, and let $L(\theta)$ be the maximal number of points one can collect along a North-East path from $(0,0)$ to (θ,θ) (as shown at Figure 11).

This quantity can be seen as a limit of the LPP times when $w_{i,j}$ takes only two values 0 and 1, and the probability of 1 is very small. Such considerations explain that $L(\theta)$ should be also in the KPZ universality class. And, indeed, this is true.

Longest Increasing subsequences in a random permutation

- The Hammersley Process can be translated to: "The number of non-intersecting broken lines that can be selected in the chosen region.

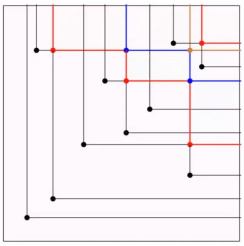


- Order the horizontal and vertical coordinates of the Poisson points and write the coordinates (x, y) of each point in the form of a biletter $\binom{x}{y}$
- The coordinates here depends only on the order of appearance and is independent of the Cartesian coordinate system and distances.
- Example:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 1 & 9 & 2 & 4 & 7 & 5 & 6 & 10 & 3 & 8 \end{pmatrix} \tag{1.7}$$

- The Hammersley Process (red line) corresponds to the longest increasing subsequence in the sequence of y coordinates.

Robinson-Schensted correspondence: geometric description



The purpose is to find a pair of Young diagrams that correspond one-to-one with the permutation (1.7).

The red points are the intersections where the black points emit rays upward and to the right, while the blue points are the focal points of the rays emitted by the red points, and so on.

Example 1.2. *The random variable ℓ_n has an interesting interpretation in terms of an airplane boarding problem. Imagine a simplified airplane with one seat in each of n rows, large distances between rows, and one entrance in front. Each entering passenger has a ticket with a seat number, but the order of passengers in the initial queue is random (this is our random permutation). Suppose that each passenger has a carry-on, and it takes one minute for that person to load it into the overhead bin as soon as (s)he reaches her/his seat. The aisle is narrow, and nobody can pass the passenger who is loading the*

carry-on. It turns out that the total time to board the airplane is precisely ℓ_n. Let us demonstrate this with an example.

Consider the permutation $\sigma = 2413$ with $\ell_\sigma = 2$. The airplane boarding looks as follows: The first passenger enters the airplane and proceeds to the seat number 2. While (s)he loads a carry-on, the other passengers stay behind and the one with the ticket for the seat number 1 ((s)he was the third person in the original queue) starts loading her/his carry-on. After one minute, the passenger with the ticket for the seat number 4 proceeds to his seat and also starts loading, as well as the one aiming for the seat number 3. In two minutes the boarding is complete.

Interestingly enough, if the queue is divided into groups, as often happens in reality, then the boarding time (for long queues) will only increase by the factor $\sqrt{k}$, where k is the number of the groups.

Let us now proceed to more recent developments. In the Last Passage Percolation problem we were maximizing a functional $H(x)$ over a set $\mathcal{X}$. A general statistical mechanics principle says that such a maximization can be seen as zero-temperature limit of the Gibbs ensemble on $\mathcal{X}$ with Hamiltonian $-H(x)$. More formally, we have the following essentially obvious statement

$$\max_{x\in\mathcal{X}} H(x) = \lim_{\beta\to\infty} \frac{1}{\beta} \ln \sum_{x\in\mathcal{X}} e^{\beta H(x)}.$$

The parameter β is usually referred to as the inverse temperature in the statistical mechanics literature.

In the Last Passage Percolation model, $\mathcal{X}$ is the set of all directed paths joining $(1,1)$ with a point (a,b), and the value of H on path x is the sum of $w_{(i,j)}$ along the path x. The Gibbs ensemble in this case is known under the name of a "Directed Polymer in Random Media". The study of such objects with various path sets and various choices of noise (i.e. $w_{(i,j)}$) is a very rich subject.

Directed Polymers in Random Media appeared for the first time close to thirty years ago in an investigation of low temperature expansion of the partition function of the Ising model with domain wall boundary conditions, but nowadays there are many other physical applications. Let us give one concrete model where such polymers arise.

Consider a set of massive particles in $\mathbb{Z}$ that evolve in discrete time as follows. At each time moment the mass of each particle is multiplied by a random variable $d_{t,x}$, where t is the time moment and x is the particle's position. Random variables $d_{t,x}$ are typically assumed to be i.i.d. Then each particle gives birth to a twin of the same mass and the twin moves to $x+1$. If we now start at time 0 with a single particle of mass 1 at $x=1$, then the mass $Z(T,x)$ of all particles at x at time T can be computed as a sum over all directed paths $(1,1) = b[1] \to b[2] \to \dots b[x+T-1] = (T,x)$ joining $(1,1)$

and (T, x):

$$Z(T,x) = \sum_{(1,1)=b[1]\to b[2]\to\ldots b[x+T-1]=(T,x)} \prod_{k=1}^{x+T-1} d_{b[k]}. \tag{1.8}$$

This model can be used as a simplified description for the migration of plankton with $d_{t,x}$ representing the state of the ocean at location x and time t which affects the speed of growth of the population. Independent $d_{t,x}$ model quickly changing media, e.g. due to the turbulent flows in the ocean.

Random Polymers in Random Media exhibit a very interesting phenomenon called intermittency which is the existence of large peeks happening with small probability, that are high enough to dominate the asymptotics of the moments. Physicists believe that intermittency is widespread in nature and, for instance, the mass distribution in the universe or a magnetogram of the sun show intermittent behavior. To see this phenomenon in our model, suppose for a moment that $d_{t,x}$ does not depend on t. Then there would be locations where the amount of plankton exponentially grows, while in other places all the plankton quickly dies, so we see very high peaks. Now it is reasonable to expect that such peaks would still be present when $d_{t,x}$ are independent both of t and x and this will cause intermittency. Proving and quantifying intermittency is, however, rather difficult.

Regarding the distribution of $Z(T, x)$, it was long believed in the physics literature that it should belong to the same KPZ universality class as the Last Passage Percolation. Now, at least in certain cases, we can prove it. The following integrable random polymer was introduced and studied by Seppäläinen [Seppäläinen-12] who proved the $t^{1/3}$ exponent for the fluctuations. The next theorem is a refinement of this result.

Theorem 1.6 (Borodin-Corwin-Remenik)**.** *Assume $d_{t,x}$ are independent positive random variables with density*

$$\frac{1}{\Gamma(\theta)} x^{-\theta-1} \exp\left(-\frac{1}{x}\right).$$

Then there exist $\theta^* > 0$ and (explicit) $c_1, c_2 > 0$ such that for $0 < \theta < \theta^*$,

$$\lim_{n\to\infty} \mathbb{P}\left(\frac{Z(n,n) - c_1 n}{c_2 n^{1/3}} \le s\right) = F_2(s).$$

The upper bound on the parameter $\theta > 0$ in this theorem is technical and it will probably be removed in future works.

1.1.9 Examples and applications

In a similar way to our transition from Last Passage Percolation to monotone paths in a Poisson field and longest increasing subsequences, we can do a limit transition here, so that discrete paths in (1.8) turn into Brownian bridges, while $d_{t,x}$ turn into the space–time white noise. Let us explain in more detail how this works as this will

provide a direct link to the Kardar–Parisi–Zhang equation that gave the name to the KPZ universality class.

For a Brownian bridge $B = B(s)$ we obtain a functional

$$H(B) = \int \beta \dot{W}(s, B(s)) ds,$$

where $\dot{W}$ is the $2d$ white noise. Thus, the partition function $Z(t,x)$ has the form $Z(t,x) = \frac{1}{\sqrt{2\pi t}} \exp\left(-\frac{x^2}{2t}\right) \mathbb{E}\left(: \exp : (H(B))\right)$, where $\mathbb{E}$ is the expectation with respect to the law of the Brownian bridge which starts at 0 at time 0 and ends at x at time t, and $: \exp :$ is the Wick ordered exponential and references therein for more details. Note that the randomness coming from the white noise is still there, and $Z(t,x)$ is a random variable.

Another way of defining $Z(t,x)$ is through the stochastic PDE it satisfies:

$$\frac{\partial}{\partial t} Z(t,x) = \frac{1}{2} \left(\frac{\partial}{\partial x}\right)^2 Z(t,x) + \dot{W} Z. \tag{1.9}$$

This is known as the stochastic heat equation. Indeed, if we remove the part with the white noise in (1.9), then we end up with the usual heat equation.

If the space (corresponding to the variable x) is discrete, then an equation similar to (1.9) is known as the parabolic Anderson model; it has been extensively studied for many years.

Note that through our approach the solution of (1.9) with δ-initial condition at time 0 is the limit of discrete $Z(t,x)$ of (1.8) and, thus, we know something about it.

If we now define U through the so-called Hopf–Cole transformation

$$Z(x,t) = \exp(U(x,t)),$$

then, as a corollary of (1.9), U formally satisfies

$$\frac{\partial}{\partial t} U(t,x) = \frac{1}{2} \left(\frac{\partial}{\partial x}\right)^2 U(t,x) + \left(\frac{\partial}{\partial x} U(t,x)\right)^2 + \dot{W}. \tag{1.10}$$

which is the non-linear Kardar–Parisi–Zhang (KPZ) equation introduced in [Kardar-Parisi-Zhang-86] as a way of understanding the growth of surfaces we started with (i.e. ballistic deposition), for a nice recent survey.

Due to the non-linearity of (1.10), it is tricky even to give a meaning to this equation, but physicists still dealt with it and that's one way how the exponent $\frac{1}{3}$ of $t^{\frac{1}{3}}$ was predicted.

If we were to characterize the aforementioned results in one phrase, we would use "integrable probability". "Integrable" here refers to explicit formulas that can be de-

rived, and also hints at parallels with integrable systems. There are direct connections, e.g. $y(s)$ defined via

$$y^2(s) = -(\ln F_2(s))'',$$

solves the (nonlinear) Painlevé II differential equation

$$y''(s) = sy(s) + 2y(s)^3.$$

Also if we define $F(x_1, \ldots, x_n; t) = \mathbb{E}(Z(x_1, t) \cdots Z(x_n, t))$, where $Z(t, x)$ is the solution of Stochastic Heat Equation (1.9), then

$$\frac{\partial}{\partial t} F = \frac{1}{2} \left(\sum_{i=1}^{n} \left(\frac{\partial}{\partial x_i} \right)^2 + \sum_{i \neq j} \delta(x_i - x_j) \right) F, \tag{1.11}$$

where δ is the Dirac delta function. (1.11) is known as the evolution equation of the quantum delta-Bose gas. It was the second quantum many body system solved via Bethe ansatz.

There is also a deeper analogy: Both integrable systems and integrable probability models can be viewed as shadows of representation theory of infinite-dimensional Lie groups and algebras. However, while integrable PDEs often represent rather exotic behavior from the point of view of general PDEs, integrable probability delivers universal behavior for the whole universality class of similar models. Moreover, in the rare occasions when the universality can be proved , one shows that the generic behavior is the same as in the integrable case. Then the integrable case provides the only known route to an explicit description of the answer.

The followings are two models in random matrices theory

Theorem 1.7 (Gaussian Unitary Ensemble)**.** *The Gaussian Unitary Ensemble (GUE) of random matrices consists of Hermitian matrices H of size $N \times N$ distributed according to the probability measure*

$$p^{GUE}(H)dH = \frac{1}{Z_N} \exp\left(-\frac{1}{2N} Tr(H^2) \right) dH,$$

where $dH = \prod_{i=1}^{N} dH_{i,i} \prod_{1 \leq i < j \leq N} dRe(H_{i,j}) dIm(H_{i,j})$ is the reference measure and Z_N the normalization constant. Denote by $\lambda_{N,max}^{GUE}$ the largest eigenvalue of a $N \times N$ GUE matrix. Then Tracy and Widom proved that the asymptotic distribution of the (properly rescaled) largest eigenvalue is F_2 :

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\lambda_{N,max}^{GUE} - 2N}{N^{1/3}} \leq s \right) = F_2(s).$$

Theorem 1.8 (Gaussian Orthogonal Ensemble)**.** *The Gaussian Orthogonal Ensemble*

(GOE) of random matrices consists of symmetric matrices H of size $N \times N$ distributed according to the probability measure

$$p^{GOE}(H)dH = \frac{1}{Z_N} \exp\left(-\frac{1}{4N} Tr(H^2)\right) dH,$$

where $dH = \prod_{1 \le i \le j \le N} dH_{i,j}$ is the reference measure and Z_N the normalization constant. Denote by $\lambda_{N,max}^{GOE}$ the largest eigenvalue of a $N \times N$ GOE matrix. The asymptotic distribution of the (properly rescaled) largest eigenvalue is F_1:

$$\lim_{n \to \infty} \mathbb{P}\left(\frac{\lambda_{N,max}^{GOE} - 2N}{N^{1/3}} \le s\right) = F_1(s).$$

1.2 Gaussian unitary ensemble: the eigenvalue point process

All information concerning the distribution of eigenvalues in any of the classical ensembles (GOE, GUE, Wishart) is encoded in the joint densities, for which we have obtained explicit formulas. Unfortunately, getting information out of the joint density requires integration, and many of the integrals that arise cannot be evaluated in any nice closed form. Nevertheless, it is possible to show that as the size N of the ensemble becomes large, certain interesting functions of the eigenvalues – for instance, the maximum – have (after suitable re-centering and re-scaling) limit distributions. The unitary ensembles are easier to handle than the orthogonal ones, so we will limit our attention to these (at least for now). To be definite, we will focus on the GUE, for which the joint distribution of the eigenvalues λ_i, listed in random order, is

$$P\{\lambda_i \in d\lambda_i \text{ for } i \le N\} = C_N \Delta_N(\chi)^2 \exp\left(-||\chi||^2/2\right) \prod_{i=1}^{N} d\chi_i = C_N \Delta_N(\chi)^2 \prod_{i=1}^{N} d\mu(\chi_i) \tag{1.12}$$

where μ is the standard normal distribution on $\mathbb{R}$ and

$$C_N = \frac{1}{(2\pi)^{N/2} N! \prod_{j=1}^{N} [j!^{-1}]}. \tag{1.13}$$

Our aim is to study the asymptotic behavior of the spectrum of certain random matrices

1.2.1 Wigner Matrices

Definition 1.1 (real Wigner matrices)**.** *For $1 \le i < j < \infty$ let $X_{i,j}$ be i.i.d. (real) random variables with mean 0 and variance 1 and set $X_{j,i} = X_{i,j}$. Let $X_{i,i}$ be i.i.d. (real) random variables (with possibly a different distribution) with mean 0 and variance 1. Then $M_n = [X_{i,j}]_{i,j=1}^n$ will be a random $n \times n$ symmetric matrix.*

Definition 1.2 (complex Wigner matrices)**.** *For $1 \leq i < j < \infty$ let $X_{i,j}$ be i.i.d. (complex) random variables with mean 0, $\mathbb{E}|X_{i,j}|^2 = 1$ and set $X_{j,i} = \overline{X_{i,j}}$. Let $X_{i,i}$ be i.i.d. (real) random variables with mean 0 and variance 1. Then $M_n = [X_{i,j}]_{i,j=1}^n$ will be a random $n \times n$ hermitian matrix.*

In both cases there are n random eigenvalues which we will denote by

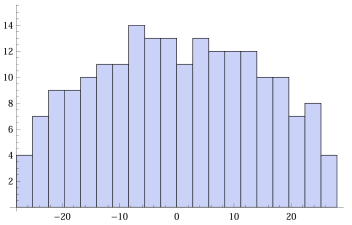
$$\lambda_1 \leq \lambda_2 \leq \dots \lambda_n.$$

(We will denote the dependence on n). Fact (which we will prove later): these are continuous functions of M_n hence they are random variables themselves.

We would like to study the scaling limit of the **empirical spectral measure**

$$\nu_n^* = \frac{1}{n} \sum_{i=1}^n \delta_{\lambda_i}.$$

This is a random discrete probability measure which puts n^{-1} mass to each (random) eigenvalue. The following picture shows the histogram of eigenvalues for a certain 200×200 Wigner matrix.



The picture suggests that there is a nice deterministic limiting behavior. In order to figure out the right scaling, we first compute the order of the empirical mean and second moment of the eigenvalues.

$$\frac{1}{n} \sum_{i=1}^n \lambda_i = \frac{1}{n} \operatorname{Tr} M_n = \frac{1}{n} \sum_{i,j=1}^n X_{i,j}$$

$$\frac{1}{n} \sum_{i=1}^n \lambda_i^2 = \frac{1}{n} \operatorname{Tr} M_n^2 = \frac{1}{n} \sum_{i,j=1}^n X_{i,j}^2$$

The first moment converges to 0 by the strong law of large numbers. However the second moment is of $O(n)$ as we have about $n^2/2$ independent terms in the sum with a normalization of $\frac{1}{n}$ instead of $\frac{1}{n^2}$. This suggests that in order to see a meaningful limit, we need to scale the eigenvalues (or the matrix) by $\frac{1}{\sqrt{n}}$.

The following theorem states that in case we indeed have a deterministic limit.

Theorem 1.9 (Wigner's semicircle law)**.** *Let*

$$\nu_n = \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i/\sqrt{n}}$$

be the normalized empirical spectral measure. Then as $n \to \infty$ we have

$$\nu_n \Rightarrow \nu \qquad a.s.$$

where ν has density

$$\frac{d\nu}{dx} = \frac{1}{2\pi}\sqrt{4-x^2}\,\mathbf{1}_{\{|x|\le 2\}}.$$

(There will be some assumptions on the distribution of the random entries of M_n, ...)

1.2.2 Gaussian Ensembles

We also discussed some special Wigner matrix models.

Definition 1.3 (GOE)**.** *Consider a real Wigner matrix where $X_{i,j} \sim N(0,1)$ and $X_{i,i} \sim \sqrt{2}N(0,1)$. The resulting random matrix model is called Gaussian Orthogonal Ensemble (or GOE). Another construction: let $a_{i,j}, i,j \in \mathbb{Z}$ be i.i.d. standard normals and $A_n = [a_{i,j}]_{i,j=1}^n$. (Note that this is not a symmetric matrix!). Then the distribution of $M_n = \frac{A_n + A_n^T}{\sqrt{2}}$ is GOE.*

It is easy to check the following useful fact: if $C \in \mathbb{R}^{n\times n}$ is orthogonal (i.e. $CC^T = I$) the $C^T M_n C$ has the same distribution as M_n. (The GOE is invariant to orthogonal conjugation.) It is a bit harder (we will prove it later) that one can actually compute the joint eigenvalue density which is given by

$$f(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_1}\prod_{i<j}|\lambda_j - \lambda_i| e^{-\frac{1}{4}\sum_{i=1}^n \lambda_i^2}.$$

Here Z_1 is an explicitly computable normalizing constant (which also depends on n).

Definition 1.4 (GUE)**.** *Consider a complex Wigner matrix where $X_{i,j}$ is standard complex Gaussian (i.e. $X_{i,j} \sim N(0,\frac{1}{2}) + iN(0,\frac{1}{2})$) and $X_{i,i} \sim N(0,1)$ (real). The resulting random hermitian matrix model is called Gaussian Unitary Ensemble (or GUE). Another construction: let $a_{i,j}, i,j \in \mathbb{Z}$ be i.i.d. standard complex Gaussians and $A_n = [a_{i,j}]_{i,j=1}^n$. (Note that this is not a symmetric matrix!). Then the distribution of $M_n = \frac{A_n + A_n^*}{\sqrt{2}}$ is GUE.*

As the name suggests, GUE is invariant under unitary conjugation. If $C \in \mathbb{C}^{n\times n}$ is unitary (i.e. $CC^* = I$) the $C^T M_n C$ has the same distribution as M_n. (We will later

show that the joint eigenvalue density is given by

$$f(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_2} \prod_{i<j} |\lambda_j - \lambda_i|^2 e^{-\frac{1}{2}\sum_{i=1}^n \lambda_i^2}.$$

Here Z_2 is an explicitly computable normalizing constant (which also depends on n).

One can see the similarity between the two densities: they are contained in the following one-parameter family of densities:

$$f(\lambda_1, \ldots, \lambda_n) = \frac{1}{Z_\beta} \prod_{i<j} |\lambda_j - \lambda_i|^\beta e^{-\frac{\beta}{4}\sum_{i=1}^n \lambda_i^2}.$$

For a given $\beta > 0$ the resulting distribution (on ordered n-tuples in $\mathbb{R}$) is called Dyson's β-ensemble. For $\beta = 1$ one gets the eigenvalue density of GOE, for $\beta = 2$ we get the GUE. The $\beta = 4$ case is also special: it is related another classical random matrix model, the Gaussian Symplectic Ensemble (GSE), which can be defined using quaternions.

For other values of β there are no 'nice' random matrices in the background. (We will see that one can still build random matrices from which we get the general β-ensemble, but they won't have such nice symmetry properties.)

Later in the semester we will show that if one scales the β ensembles properly ('zooming in' to see the individual eigenvalues near a point) then one gets a point process limit. The limiting point process is especially nice in the $\beta = 2$ case (GUE). It is conjectured that its distribution appears among the critical line zeros of the Riemann-ζ function.

References

[1] Alexei Borodin and Leonid Petrov. "Integrable probability: From representation theory to Macdonald processes". In: (2014).

[2] Alexei Borodin and Vadim Gorin. "Lectures on integrable probability". In: *Probability and statistical physics in St. Petersburg* 91 (2016), pp. 155–214.

[3] Alexei Borodin and Leonid Petrov. "Integrable probability: stochastic vertex models and symmetric functions". In: *Stochastic processes and random matrices*. Oxford University Press Oxford. 2017, pp. 26–131.

[4] Patrik L Ferrari. "Introduction to Integrable Probability". In: (2020).