Gergen Lectures, Duke University

October 22-24, 2001

Random Matrix Theory

Lecture I

It is a great pleasure and honor for me to give the year 2001 Gergen Lectures. I am also very pleased to be here with so many of my friends. My topic is random matrix theory with emphasis on the relationship to integrable systems.

In this first part of a lecture, I am going to begin with some general remarks about integrable systems. Then I will make some general and historical remarks about random matrix theory (RMT).
• Universality conjecture in RMT

In the 2nd lecture, I will

• introduce a variety of objects and techniques from modern theory of integrable systems

In the 3rd lecture, I will show how

• to use these techniques to solve the universality conjecture following the approach of T. Kriecherbauer, K. McLaughlin, S. Venakides, X. Zhou and P.D.

In the 4th and final lecture, I will show how

• to use these techniques to analyze the so-called Ulam problem in the theory of random permutations.

Integrable systems

The modern theory of integrable systems began with the discovery of Gardner, Greene, Kruskal and Miura of a method for integrating the Korteweg de Vries (KdV) equation

\[ u_t + uu_x + u_{xxx} = 0, \quad x \in \mathbb{R}, \quad t \geq 0 \]

\[ u(x, t=0) \to 0 \quad \text{as} \quad |x| \to \infty \]

This equation arises in the theory of water waves.
Initially, the discovery of Gardner et al. was regarded as providing a method of solution for a rather thin set of evolutionary equations. But by the early 1980s, it started to become clear that the discovery of Gardner et al. was just the first glimpse of a far more general integrable method. That would eventually have applications across the broad spectrum of problems in pure and applied mathematics. In the narrowest sense, an integrable system is a
Hamiltonian dynamical system — finite or infinite dimensional — with "enough" integrals of the motion, all of whose Poisson brackets are zero, to reduce the system in some "explicit" form. Now it has been a rather extraordinary experience in the field over the last 30 years, that many systems which are of great mathematical and physical interest, which may be Hamiltonian, and may not even be dynamical, can be solved "explicitly" using techniques that have a direct link back to the method of solution for the KdV equation discovered by Gardner et al. A kind of developments that I have in mind are for example

- The resolution of the classical Schottky problem in algebraic geometry in terms of the solution of
* The introduction of quantum groups

* integrable statistical models — connection to Jones polynome

* 2D quantum gravity and the work of Witten and Kon on the KP hierarchy

* nonlinear special function theory (Painlevé Theory)

* conformal field $\mathfrak{g}^\ast$ — work of Kac, Dobrev

* random matrix theory

* combinatorial problems of Robinson-Schensted-Knuth

The thrust of these lectures is to illustrate how

the integrable methods interact with one of these topics

— random matrix theory.

Allow me a very schematic moment. Consider

the following picture

![Diagram](image-url)
In 1967, these two "manifolds" had a "transverse intersection". The initial thought was that to draw the ideas of Gardner et al. one should move along the "pole" manifold. But this turned out to be too limiting; we now know that in order to describe the full development of the method one should move in the "transverse" direction. And this in the direction we will move in the next 2 lectures.

RMTR.

Now what in random matrix theory? RMTR was first studied in the 1930s in the context of mathematical statistics by Hsu-Wishart and others, but it was Wigner in the early 1950s who introduced RMTR to mathematical physics. Wigner was concerned with the scattering of neutrons off large...
Schematically one has the following:

The energies $E_i$ for which one obtains a large scatter cross section are called resonances, and for a heavy nucleus like $U^{238}$, say, there can be hundreds of them. In theory, one could write down the Schrödinger equation for the neutrons + nucleus and try to solve it numerically to compute these $E_i$. But in Wigner's time, and still in our time...
also into the forseeable future, this was not a real approach, and so people began to think that it was more appropriate to give the resonances a statistical meaning. But what should the statistical model be? At this point Wigner put forward the remarkable hypothesis that the (high) resonances $E_i$ behave like the eigenvalues of a (large) matrix. It is to overemphasize what a radical and revolutionary this was: we all recall that when we first were learning some physics, we understood that the detail of the model were paramount; if you changed the force law in the equations of motion, the behavior of the system would change. But now all of that was out the window. The precise mechanism was
no longer important. All that Wigner (required was the
real symmetric)
• the matrices be Hermitian (so that the eigenvalues were real)

• that the ensemble behaves "naturally" under certain
physical symmetry groups (GUE, GOE, GSE ensemble — more later)

During the late 50's, 60's and early 70's, various research
workers began to test Wigner's hypothesis against real experimen-
data in a variety of physical situations and the
results were pretty impressive. A classic reference for
RMT is

Random matrices — M. L. Mehta, 2nd Ed

and I invite you to look at p17 for a comp
of 1726 nuclear level spacings against predictions of
the GOE ensemble.
procedure: first we must rescale the resonances and the eigenvalues of the random matrices so that

\[
\text{number of resonances/unit interval} = \frac{\text{expected number of eigenvalues/unit interval}}{n}
\]

Then we compare the statistics.

More concretely, we typically have a situation where we are interested in resonances in an interval \((E, E + \delta E)\), where \(\delta E \ll E\) but \(\delta E\) contains many resonances.

We then rescale the resonances in \(\delta E\) so that their average density is one. We then do a similar scaling for the eigenvalues in some ensemble, as only at that point, after we have adjusted our "microscopes," do we compare statistics. A scientist approaches such problems with two instruments in — a microscope and a list of ensembles.
The dice on the microscope are adjusted to the matrix of the situation, and vary from system to system.

But once the "slide" is in focus, one sees universal behavior described by one of the entries in the list of ensembles.

Henry gave me a schematic moment. I now ask you for a moment when began to think of something that is quite deterministic in a statistical way. Nevertheless, is a process with which we are very familiar. I take, for example, a die, which surely obeys the laws of solid mechanics, which we readily and intuitively understand as a statistical object. Moreover, we "know" what the stochastic model should be: all 6 sides have equal probability. Looked at from this point of view, what Wigner had to do was to see the neutron +...
it's stochasticity. In all such problems there is a
some singular/asymptotic process) and when we cross some "Bayesian
point, phase space opens up, and all bets are off.

We are standing, as it were, on the corner of 7th
and 6th Street and we are watching this little kid
play 3 card brag: if we are fast enough, we
can follow all his moves but then "poof!" The cards
are out there, and all 3 are equally likely.

---

Now in the early 1970s a very remarkable
thing happened. Montgomery, quite independently of the
other going on, began thinking that the zeros of the
Riemann zeta function on the critical line \( \zeta = \frac{1}{2} \)
should also be viewed statistically. And so, assume

\( \zeta = \frac{1}{2} \)
\[
\left( \frac{1}{N} \sum_{\text{pairs } (i, \delta)} \prod \right) \text{in an interval } \left( E, E + \delta E \right) \subset E > 1,
\]

\[ E + \delta E = \bar{E}, \quad \tilde{s}_i = \bar{s}_i + i \tilde{\delta}_i \]

\[ \text{rescaled: } \tilde{\delta}_i = \frac{\bar{s}_i - \log \bar{s}_i}{\log \bar{T}} \]

and he found a limiting formula \( R = R(\alpha, \beta) \)

\[
\left( \int_{\alpha}^{\beta} \left[ 1 - \left( \frac{\sin \pi u}{\pi u} \right)^2 \right] du \right)
\]

for the 2 pt. function as \( E \to \infty, \# \delta_i \to \infty \). What happened next is very well-known (I even asked Dyson to authenticate this version). Montgomery met Dyson at the Institute in Princeton, and when he told him about his calculations, Dyson immediately wrote down a formula and asked Montgomery, "Did you get this"
show he knew the answer, Dyson said, "Well, if we go of the zeros function behaved like the eigenvalues of a random GUE matrix, this would have to be the answer!" Indeed, what Montgomery had obtained for $R$ was precisely the 2-point function for the eigenvalues of a (large) random (GUE).

At this point, the cat was out of the bag. People began asking whether their favorite list of the zeros behaved like the eigenvalues of a random matrix.

Through the 80s, an extraordinary variety of systems were investigated from this point of view with astounding regularity: for example, if we take a region in the plane
and look at the eigenvalues of the Dirichlet Laplacian in this region. They too (after the standard scaling) behavior like the eigenvalues of a random GUE matrix (see Mehta p13).

In the late 1980’s, Montgomery’s work was up numerically by Odlyzko, who then confirmed Montgomery’s work to high accuracy and also investigated other statistics such as the nearest neighbor spacing again there was incredible agreement with random matrix theory. In recent years, the work of Montgomery and Odlyzko has been a wonderful springboard for Sarnak–Rudnick and Ten–Sarnak–Katz, to prove all kinds of GUE (and GSE) random matrix properties, including the zeros of all kinds of automorphic L functions.
Now up until very recently, the physical and mathematical phenomena which were investigated concerned the eigenvalues of a random matrix in the bulk of the spectrum.

But in the last 2 yrs or so a very interesting class of problems have started to appear in combina and also in Statistical particle models, which concern the eigenvalues at the top of the spectrum. For example consider the following version of solitaire called "patience sorting". Suppose we have $N$ cards numbered linearly 1, 2, ..., $N$ for convenience.

Shuffle the cards

1 2 3 ... $N$
Now take the top card \( \pi_1 \) ...  

Eventually you end up with \( P_n(\pi) \). Question

Put any uniform distribution on \( \pi \). Do \( P_n(\pi) \) behave as \( n \to \infty \)? Or more generally, when do I need typically to play the game with \( n \) cards?

To state the answer to this question we need to introduce some notation which we will discuss in the next lecture, the theorem is the following (Baik, Johansson, P.D.)

As \( n \to \infty \), \( P_n \), suitably centered and scaled, behaves statistically like the largest eigenvalue of a random GUE matrix.

More precisely, assume \( X^{(n)}(m) \sim X^{(n)}(1) \cdots X^{(n)}(m) \).
The eigenvalues of a (large) $N \times N$ GUE matrix $V$ satisfy

$$\lim_{N \to \infty} \text{Prob} \left( \frac{L_n - 2\sqrt{N}}{N^{1/6}} \leq t \right) = \lim_{N \to \infty} \text{Prob} \left( \frac{L_n - 2\sqrt{N}}{2^{1/6} N^{1/6}} \right) = F(t)$$

$F(t)$ is called the Tracy-Widom distribution, after its explicit discovery and can be expressed in terms of a solution of the Painlevé II equation.

There are many equivalent formulations of the patience waiting problem — long-standing results follow from a random permutation. The height of a random directed acyclic graph model, the number of boxes in the first row of a Young diagram under Plancherel measure.
(c) \[ P_{(m)} = \frac{1}{m!} (e^{-\lambda I}) P_{(m)} (e^{-\lambda I}) \] which in invariant under non-mutable Hamiltonian variation \( P_{(m)} = P_{(m)} \) since \( T \parallel \hat{e}_a \parallel T^\dagger \hat{e}_a \).

(a) Gaussian Unitary Ensemble (GUE)

(b) \[ N \times N \]

consisting of

\[ P_{(m)} \] Wigner matrix

\[ P_{(m)} \] is the

determinant

of the

behavior

of the

system

under


time

reversal

...  

What does this mean? Statistically, like the 2nd larger eigenvalue. 

New in this paper...
2) Gaussian Orthogonal ensemble (GOE)
   (a) N x N real symmetric

3) Gaussian Symplectic ensemble (GSE)
   (a) 2N x 2N Hermitian self dual
       \[ M = M^* \]
       \[ M = J M^T J \]
       \[ J = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix} \]
   (b) 
   Invariant under \( U^* M U \)
       Symplectic
       \[ U^T J U = J \]

The analysis of GUE is the simplest and I am going to restrict myself to this case throughout these lectures.

So focusing on GUE, the first theorem is the linearity which is that if \( P(U) \) satisfies (a)-(b)-(c) then necessarily

\[ P(U) = \text{const} \ e^{-\Delta(K^2 + \beta n + \delta)} \]

with \( \Delta, \beta, \gamma \in \mathbb{R} \). Centering and rescaling we have the GUE distribution

\[ P(U) \propto e^{-\frac{1}{2} K^2} \ dv \]

(2)
Now here comes the problem: whereas conditions (a) and (b) are physical, (c) is just a device and has no physical basis.

If we just assume (a) and (b) we find

\[ P(W|\lambda) = \frac{1}{Z} e^{-\lambda V(W)} \, dW \]

for some real valued function $V$, $V(x) \to \infty$ as $|x| \to \infty$ (called Unitary Ensembles).

Ensembles of matrices of type (3) are called Unitary Ensembles. Now physicists turn all on its head and say, "There is no physical way to distinguish between different choices of $V$. Then to physical questions such as whatever answers we compute, the answer must be independent of $V". This is a rough form of what is meant by universality in random matrix theory. In the coming lectures we will focus on a particular class.
The probability that a (large) matrix has no eigenvalues in an interval \((a,b)\), and we will show that \(P(a,b)\) indeed has a universal form independent of \(V\). And in proving this, we will employ techniques that have arisen in integrable Theory over the years and can be traced back in some form to the original method of solution for the KdV equation discovered in 1967 by Gardner, Greene, Kruskal and Miura.

To summarize, what has emerged, somewhat mysteriously and to everyone's surprise is a very powerful heuristic: all kinds of things, physical and mathematical, behave, in some limit, like random matrices. Next to you have a system that you want to approximate from a large matrix, so on and so forth. You have a good chance of be