Eigenvalue Distribution of Large N Random Matrix—A Diagrammatic Proof to Marchenko Pastur Law

Xiaochuan Lu[1,2,*] and Hitoshi Murayama[1,2,3,†]

[1]*Department of Physics, University of California, Berkeley, California 94720, USA*
[2]*Theoretical Physics Group, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*
[3]*Kavli Institute for the Physics and Mathematics of the Universe (WPI), Todai Institutes for Advanced Study, University of Tokyo, Kashiwa 277-8583, Japan*

The large N limit is a very useful tool in various theoretical models. In random matrix theory, Marchenko-Pastur law is a powerful theorem that governs eigenvalue distribution of large dimension random matrices. We present an alternative proof of Marchenko Pastur law using Feynman diagrams.

I. INTRODUCTION

In various theoretical models, large N limit has been a very useful tool, both at a qualitative and a quantitative level [1]. It works as a great approximation, even when the actual dimension of the matrix is not a very large number, such as $N_c = 3$ in QCD [2, 3] and $N_f = 3$ in neutrino anarchy [4]. It is thus useful to understand the behavior of large N random matrices as much as possible.

In random matrix theory, there is a powerful law—Marchenko-Pastur (MP) law [5]—that governs the asymptotic eigenvalue distribution of large-dimension random matrices. In this paper we provide an alternative proof to it.

We first state Marchenko Pastur law in Section II. Then in Section III, we present a detailed proof of MP law using Feynman diagrams. We extend our proof to symmetric matrix case in Section IV. Section V is our conclusion.

II. MARCHENKO PASTUR LAW

After the original paper [5], a lot of work followed and the theorem is sharpened and extended to a few different versions [6]. In this paper we focus on its following version, which is most related to large N models in physics.

Let X be a $M \times N$ random complex matrix, whose entries X_{ij} are generated according to the following conditions:

$$\text{(1) independent, identical distribution (i.i.d.)}, \tag{1}$$

$$\text{(2) } \langle X_{ij}\rangle_X = 0, \langle X_{ij}^2\rangle_X = 0, \text{ and } \left\langle |X_{ij}|^2\right\rangle_X = 1, \tag{2}$$

$$\text{(3) } \left\langle |X_{ij}|^{2+\varepsilon}\right\rangle_X < \infty \text{ for any } \varepsilon > 0 \tag{3}$$

where and throughout this paper, we use $\langle \mathcal{O}\rangle_X$ to denote the expectation value of a random variable $\mathcal{O}$ under the ensemble of X. Then construct an $M \times M$ hermitian matrix $A = \frac{1}{N}XX^\dagger$, whose eigenvalues are denoted by λ_k, with $k = 1, 2, ..., M$. Then the empirical distribution of these eigenvalues is defined as

$$F_M(x) \equiv \frac{1}{M}\sum_{k=1}^{M} I_{\{\lambda_k \le x\}}, \tag{4}$$

where I_B denotes the indicator of an event B:

$$I_{\{B\}} = \begin{cases} 1 & \text{if } B \text{ is true} \\ 0 & \text{if } B \text{ is false} \end{cases} \tag{5}$$

Consider the limit $N \to \infty$. If the limit of the ratio M/N is finite

$$b \equiv \lim_{N\to\infty} M/N \in (0, \infty), \tag{6}$$

then $\langle F_M(x)\rangle_X \to F(x)$, where $F(x)$ denotes the cumulative distribution function of the Marchenko-Pastur distribution whose density function is

$$f(x) = \frac{1}{2\pi}\frac{\sqrt{(x_2 - x)(x - x_1)}}{x}\frac{1}{b}\cdot I_{\{x\in(x_1,x_2)\}} + (1 - \frac{1}{b})\delta(x)\cdot I_{\{b\in[1,\infty)\}}\ , \tag{7}$$

with $x_1 = (1 - \sqrt{b})^2$ and $x_2 = (1 + \sqrt{b})^2$. In the special case of a square matrix X, namely $b = 1$, this becomes

$$f(x) = \frac{1}{2\pi}\sqrt{\frac{4}{x} - 1}\cdot I_{x\in(0,4)}. \tag{8}$$

In some theoretical models of physics, the large N random matrix in concern has to be symmetric, with only $N(N+1)/2$ i.i.d. entries, such as the Majorana mass matrix in neutrino anarchy [7–9]. Using our alternative Feynman diagram approach, we also prove that MP law still holds for symmetric matrix X (in Sec. IV).

III. PROOF OF MARCHENKO PASTUR LAW WITH FEYNMAN DIAGRAMS

A. Stieltjes Transformation

For a single matrix X generated, the distribution density of eigenvalues is

$$\rho_X(E) = \frac{1}{M}\sum_{k=1}^{M}\delta(E - \lambda_k). \tag{9}$$

* luxiaochuan123456@berkeley.edu
† hitoshi@berkeley.edu, hitoshi.murayama@ipmu.jp

Our goal is to compute its expectation

$$\rho(E) \equiv \langle \rho_X(E) \rangle_X = \int dX \cdot \rho_X(E), \tag{10}$$

and prove that $\rho(E)$ approaches the MP density function (Eq. 7) as $N \to \infty$. Here we use dX to denote the normalized measure of X:

$$dX = \prod_{ij} g(X_{ij}) dX_{ij}, \quad \int dX = 1, \tag{11}$$

with $g(X_{ij})$ denoting the normalized distribution density of each X_{ij}.

Since $\rho(E)$ is not easy to compute directly, we make use of a method known in mathematics as "Stieltjes transformation". That is, from the identity, where x is a real variable

$$\delta(x) = -\frac{1}{\pi} \lim_{\varepsilon \to 0^+} \mathrm{Im} \frac{1}{x + i\varepsilon}, \tag{12}$$

we get

$$\begin{aligned} \rho(E) &= \int \rho(x) \delta(E - x) dx \\ &= -\frac{1}{\pi} \lim_{\varepsilon \to 0^+} \mathrm{Im} \int \frac{\rho(x)}{E + i\varepsilon - x} dx. \end{aligned} \tag{13}$$

Thus for any distribution density function $\rho(x)$, we can define its Stieltjes transformation $G(z)$, a complex function as an integral over the support of $\rho(x)$

$$G(z) \equiv \int \frac{\rho(x)}{z - x} dx. \tag{14}$$

Then according to Eq. (13), $\rho(x)$ can be obtained from the inverse formula

$$\rho(E) = -\frac{1}{\pi} \lim_{\varepsilon \to 0^+} \mathrm{Im} G(E + i\varepsilon). \tag{15}$$

For our case, $G(z)$ can be computed as following

$$\begin{aligned} G(z) &= \int \frac{\rho(x)}{z - x} dx = \int \frac{1}{z - x} \langle \rho_X(x) \rangle_X dx \\ &= \left\langle \frac{1}{M} \sum_{k=1}^{M} \int \frac{1}{z - x} \delta(x - \lambda_k) dx \right\rangle_X \\ &= \left\langle \frac{1}{M} \sum_{k=1}^{M} \frac{1}{z - \lambda_k} \right\rangle_X \\ &= \left\langle \frac{1}{M} tr(\frac{1}{z - A}) \right\rangle_X = \frac{1}{M} \frac{1}{z} tr\left[B(z)\right], \end{aligned} \tag{16}$$

where we have defined a matrix $B(z)$ as

$$B(z) \equiv \left\langle \frac{z}{z - A} \right\rangle_X \tag{17}$$

$$= \left\langle \sum_{n=0}^{\infty} (\frac{A}{z})^n \right\rangle_X = \left\langle \sum_{n=0}^{\infty} (\frac{1}{zN} X X^\dagger)^n \right\rangle_X. \tag{18}$$

This expansion is a valid analytical form of $B(z)$ in the vicinity of $z = \infty$. We will compute $B(z)$ in this vicinity first, and then analytically continue it to the whole complex plane. Once $B(z)$ is obtained, $G(z)$ and $\rho(E)$ would follow immediately. The following several subsections are devoted to calculate this *target function* $B(z)$ under the limit $N \to \infty$.

B. Group into "Boxes"

Our target function is a sum of various terms, in which a typical n-term looks like

$$\begin{aligned} B_{ij}(z) \supset (\frac{1}{zN})^n \left\langle \prod_{p=1}^{n} X_{\alpha_p \beta_p} X^\dagger_{\beta_p \alpha_{p+1}} \right\rangle_X \\ = (\frac{1}{zN})^n \left\langle X_{i\beta_1} X^\dagger_{\beta_1 \alpha_2} \cdots X_{\alpha_n \beta_n} X^\dagger_{\beta_n j} \right\rangle_X, \end{aligned} \tag{19}$$

with an identification $\alpha_1 \equiv i, \alpha_{n+1} \equiv j$ and a sum over all the dummy indices $\alpha_2, \alpha_3, \cdots, \alpha_n$ from 1 to M, and $\beta_1, \beta_2, \cdots, \beta_n$ from 1 to N.

As stated in the condition Eq. (1), different elements X_{ij} are independent. Therefore any such n-term expectation can be factorized

$$\begin{aligned} &\left\langle X_{i\beta_1} X^\dagger_{\beta_1 \alpha_2} \cdots X_{\alpha_n \beta_n} X^\dagger_{\beta_n j} \right\rangle_X \\ &= \langle f(X_{k_1 l_1}) \rangle_X \langle f(X_{k_2 l_2}) \rangle_X \cdots, \end{aligned} \tag{20}$$

where each individual expectation $\langle f(X_{kl}) \rangle_X$ contains only X_{kl} and its complex conjugate X^*_{kl}

$$\langle f(X_{kl}) \rangle_X = \langle (X_{kl})^{m_1} (X^*_{kl})^{m_2} \rangle_X. \tag{21}$$

Namely that the independence among elements allows us to group same X_{kl} (and the complex conjugate) together into one factor. For future convenience, let us call each such factor a "box".

Now in evaluating the n-term (Eq. 19), the sum over the dummy indices α's and β's can be decomposed into two steps. (1) There are many ways to group the $2n$ elements into boxes. We need to sum over all possible grouping configurations. (2) Under each grouping configuration, all α's within the same box are forced equal, so are all β's, thus some dummy indices are tied to others and hence no longer free to sum. But generically there are still some *free dummy indices* remaining, which needs to be summed. In summary, this decomposition of sum can be expressed as

$$\sum_{\alpha,\beta} = \sum_{\text{grouping configurations}} \sum_{\text{free dummy indices}}. \tag{22}$$

C. Key Statement in Power Counting of N

To evaluate an n-term (Eq. 19) under $N \to \infty$, an efficient way to count the power of N is definitely crucial. The suppression factor $\frac{1}{N^n}$ in front of Eq. (19) contributes a factor N^{-n}. On the other hand, with the sum decomposition Eq. 22, under each grouping configuration, summing

over each free dummy index will give us one power of N: $\sum_{\alpha=1}^{M} \delta_{\alpha\alpha} = M = bN$, $\sum_{\beta=1}^{N} \delta_{\beta\beta} = N$. From this competition, we end up with a factor $N^{-(n-n_f)}$, where n_f denotes the total number of free dummy indices under a given grouping configuration. There are $2n-1$ dummy indices in total: $\alpha_2, \cdots, \alpha_n$, $\beta_1, \cdots, \beta_n$, but not all of them are free, because within each box, the different α's and β's are forced into same value respectively. Apparently, n_f largely depends on the grouping configuration.

To study how n_f depends on the grouping configuration, we resort to some graphical analysis. First, let us draw a map to represent each grouping configuration, where each grouping box is drawn as an isolated island. We notice that in the sequence of Eq. (19)

$$X_{i\beta_1} X^\dagger_{\beta_1\alpha_2} \cdots X_{\alpha_n\beta_n} X^\dagger_{\beta_n j}, \tag{23}$$

every dummy index appears twice, i.e. in pair. Each such dummy index pair could be grouped into the same box, or two different boxes. If any two boxes share a dummy index pair, let us connect those two islands by a "bridge" on the map. So every grouping configuration is described by a map of boxes (or islands) and a number of bridges connecting them.

Suppose that there are m boxes fully connected by bridges. Within each box, after all α's and β's are identified respectively, we are left with at most 2 free dummy indices. Then collecting all the m boxes, we get at most $2m$ free indices. But to connect all these m boxes, we need at lease $m-1$ bridges. If we remove all the redundant bridges and thus chop the map into a tree map, then each of the remaining $m-1$ bridges would effectively reduce one free index out of the $2m$. Thus we have arrived at our *key statement*:

If m boxes are fully connected by bridges, then we can get at most $m+1$ free dummy indices out of them.

Clearly, the whole sequence Eq. (23) is fully connected by dummy-index bridges, so we can apply this key statement to it. Suppose there are n_b boxes in total for a grouping configuration, then we get $n_f \leq n_b + 1$. However, this upper bound is obtained by counting i and j also as dummy indices. But they are not. So we need to further subtract 1, if i and j are grouped into the same box; or subtract 2 if they are not. To sum up, we get

$$n_f \leq \begin{cases} n_b & i, j \text{ in same box} \\ n_b - 1 & i, j \text{ not in same box} \end{cases} \tag{24}$$

D. Association with Feynman Diagrams

With the result Eq. (24), we are ready to study what kinds of grouping configurations can give nonzero contribution. Due to the condition Eq. (2), each box needs to contain at least two elements in order not to vanish. But there are in total only $2n$ elements in an n-term. So if any box has more than two elements, then the total number of boxes n_b must be less than n. Consequently, $n_f \leq n_b < n$ and the grouping configuration is suppressed by a factor $N^{-(n-n_f)}$. Due to condition Eq. (3), the coefficient multiplying this factor must be finite. Thus the contribution from this kind of grouping configuration vanishes under $N \to \infty$. To sum up, only grouping the elements by pairs can give nonzero contributions.

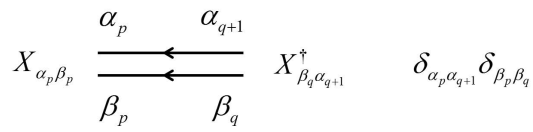


FIG. 1. Feynman rule propagator

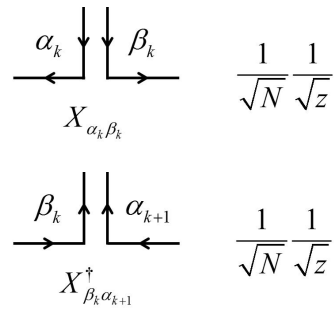


FIG. 2. Feynman rule vertices

We can call each such grouped pair a "contraction". According to the condition Eq. (2), only contracting X and $X^\dagger$ can be nonzero

$$\left\langle X_{ij} X^\dagger_{kl} \right\rangle_X = \langle X_{ij} X^*_{lk} \rangle_X = \delta_{il}\delta_{jk}. \tag{25}$$

Let us call this contraction "propagator", which corresponds to the Feynman rule shown in Fig. 1.

Here is also a bonus result: Even under pair grouping configurations where $n_b = n$, i and j must be grouped into the same box in order to get large enough n_f (see Eq. 24). So i and j must be equal, which means that the matrix $B(z)$ must be diagonal under $N \to \infty$.

Our goal is to evaluate the target function $B(z)$ (Eq. 18). Since each matrix element has two indices, and each pair of $XX^\dagger$ always comes with a factor $\frac{1}{zN}$, we are naturally led to the Feynman rules of vertices shown in Fig. 2. The arrow flow is used to distinguish X from $X^\dagger$. This is necessary because our X is generically not hermitian. Then a typical n-term (Eq. 19) can be calculated by summing over Feynman diagrams corresponding to all the possible contraction structures of Fig. 3. To give a few examples, we enumerate all nonzero diagrams contributing to $n = 0$, $n = 1$ and $n = 2$ terms in Fig. 4.

E. Simplification: Planar Diagrams only for $N \to \infty$

Now we have developed a diagrammatic way of evaluating $B_{ij}(z)$ as described by Fig. 3, which is well organized and

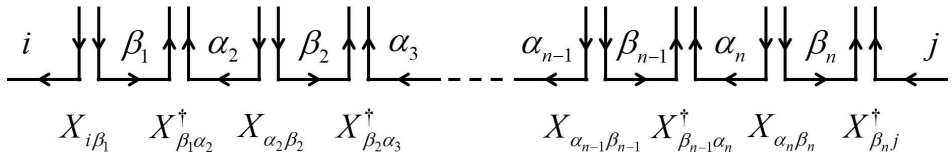
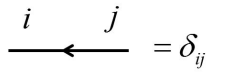
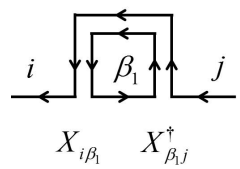


FIG. 3. Feynman diagrams of n-term (before specifying contraction structure).



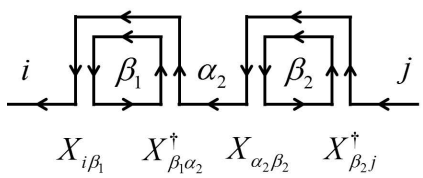
$= \delta_{ij}$

(a) $n = 0$ term



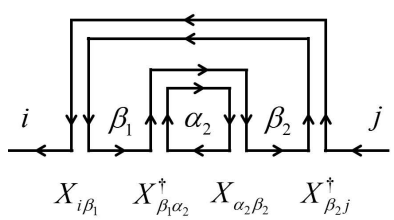
$$= \sum_{\beta_1=1}^{N} \frac{1}{zN} \delta_{ij} \delta_{\beta_1\beta_1} = \frac{1}{z} \delta_{ij}$$

(b) $n = 1$ term



$$= \sum_{\beta_1,\beta_2=1}^{N} \sum_{\alpha_2=1}^{M} \frac{1}{z^2 N^2} \delta_{i\alpha_2} \delta_{\beta_1\beta_1} \delta_{\alpha_2 j} \delta_{\beta_2\beta_2} = \frac{1}{z^2} \delta_{ij}$$

(c) $n = 2$ case 1



$$= \sum_{\beta_1,\beta_2=1}^{N} \sum_{\alpha_2=1}^{M} \frac{1}{z^2 N^2} \delta_{ij} \delta_{\beta_1\beta_2} \delta_{\alpha_2\alpha_2} \delta_{\beta_1\beta_2} = \frac{b}{z^2} \delta_{ij}$$

(d) $n = 2$ case 2

FIG. 4. Feynman diagram examples.

quite routine. But the actual calculation is still rather complicated, because there are so many ways of contracting the vertices. Large N limit, however, brings us another great simplification: Any diagram with crossed contractions will vanish under $N \to \infty$. This means that we only need to consider the type of contraction shown in the first line of the following, but not that kind shown in the second line.

$$\cdots X_{\alpha_2\beta_2} X^\dagger_{\beta_2\alpha_3} X_{\alpha_3\beta_3} X^\dagger_{\beta_3\alpha_4} X_{\alpha_4\beta_4} X^\dagger_{\beta_4\alpha_5} \cdots$$

$$\cdots X_{\alpha_2\beta_2} X^\dagger_{\beta_2\alpha_3} X_{\alpha_3\beta_3} X^\dagger_{\beta_3\alpha_4} X_{\alpha_4\beta_4} X^\dagger_{\beta_4\alpha_5} \cdots$$

Once crossed contractions are forbidden, all the propagators can only form two types of structures: "side by side" as in the example of Fig. 4(c) or "nesting" as in Fig. 4(d). A combination of these two types gives us a general "planar" diagram. Only planar diagrams have nonzero contributions under $N \to \infty$.

This requirement also follows from our key statement. Assume that we have a contraction jumping k couples of elements:

$$\cdots X^\dagger_{\beta_{p-1}\alpha_p} X_{\alpha_p\beta_p} X^\dagger_{\beta_p\alpha_{p+1}} \cdots X_{\alpha_q\beta_q} X^\dagger_{\beta_q\alpha_{q+1}} X_{\alpha_{q+1}\beta_{q+1}} \cdots$$

with $q = p + k$. This contraction identifies α_{q+1} with α_p and β_q with β_p. After summing over these non-free dummy indices α_{q+1} and β_q, we get the result proportional to (with a finite coefficient)

$$\cdots X^\dagger_{\beta_{p-1}\alpha_p} \cdot X^\dagger_{\beta_p\alpha_{p+1}} \cdots X_{\alpha_q\beta_p} \cdot X_{\alpha_p\beta_{q+1}} \cdots \tag{26}$$

Now we are only left with $n-1$ couples of X and $X^\dagger$. With the number of boxes $n_b = n - 1$ then, it seems impossible to make $n_f = n$, according to our previous result Eq. (24). However, by a careful look at the new sequence Eq. (26), we realize that it is no longer guaranteed to be fully connected by bridges. Instead, it consists of two parts: inside the contraction and outside the contraction, each part fully connected. So

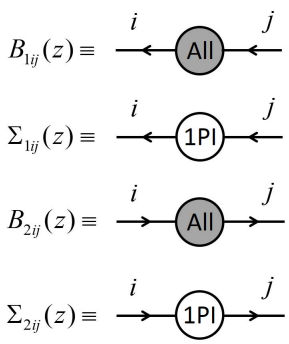


FIG. 5. Auxiliary functions defined in terms of planar Feynman diagrams.

as long as we do not group any element inside with any element outside into one box (i.e. no crossed contraction!), we can only apply our key statement to each part separately. In this case, we are just lucky enough to save it: we can get up to $k+1$ free dummy indices from the inside part and $n-1-k$ from the outside part. Together, we can still make $n_f = n$. On the other hand, if we do make a contraction crossed with the first one, then the divided two parts are reconnected through this contraction box and Eq. (24) can be applied to the whole sequence Eq. (26): $n_f \leq n_b = n-1 < n$. Therefore diagram with crossed contractions will vanish under $N \to \infty$.

F. Diagrammatic Calculation for $N \to \infty$

Now we are finally ready to calculate our target function $B_{ij}(z)$ by summing over all the planar Feynman diagrams formed from Fig. 3. For convenience, let us define four functions as shown in Fig. 5, two 1PI (1 Particle Irreducible) functions $\Sigma_{1ij}(z)$, $\Sigma_{2ij}(z)$, and two two-point functions $B_{1ij}(z)$, $B_{2ij}(z)$. Here B_{1ij} is nothing but our target function $B_{ij}(z) = B_{1ij}(z)$.

First, let us study the 1PI functions. $\Sigma_{1ij}(z)$ sums over all the 1PI planar diagrams with external single arrows pointing to the left. Each 1PI planar diagram must have a double-line contraction coating it at the most outside, with nested inside anything. Clearly, the sum of the nested part gives nothing but $B_2(z)$. So we get a relation as shown in Fig. 6(a):

$$\begin{aligned}\Sigma_{1ij} &= \sum_{\beta_p,\beta_q=1}^{N} \frac{1}{zN}\delta_{ij}\delta_{\beta_p\beta_q}B_{2\beta_p\beta_q} \\ &= \left(\frac{1}{zN}\sum_{\beta_p=1}^{N} B_{2\beta_p\beta_p}\right)\delta_{ij} \equiv \Sigma_1\delta_{ij}. \end{aligned} \tag{27}$$

We see that $\Sigma_{1ij}(z)$ is proportional to the identity matrix. There is a similar relation for $\Sigma_{2ij}(z)$ (as shown in Fig. 6(b)), which is also proportional to identity matrix:

$$\begin{aligned}\Sigma_{2ij} &= \sum_{\alpha_p,\alpha_q=1}^{M} \frac{1}{zN}\delta_{ij}\delta_{\alpha_p\alpha_q}B_{1\alpha_p\alpha_q} \\ &= \left(\frac{1}{zN}\sum_{\alpha_p=1}^{M} B_{1\alpha_p\alpha_p}\right)\delta_{ij} \equiv \Sigma_2\delta_{ij}. \end{aligned} \tag{28}$$

Now let us turn to the two point functions $B_{1ij}(z)$ and $B_{2ij}(z)$. Same as in computing a two-point correlation function in QFT, all the diagrams contributing to $B_{1ij}(z)$ ($B_{2ij}(z)$) can be organized into a geometric series of the 1PI functions $\Sigma_{1ij}(z)$ ($\Sigma_{2ij}(z)$). Since both $\Sigma_{1ij}(z)$ and $\Sigma_{2ij}(z)$ are proportional to identity matrix, $B_{1ij}(z)$ and $B_{2ij}(z)$ are also proportional to identity matrix:

$$\begin{aligned} B_{1ij} &= \delta_{ij} + \Sigma_{1ij} + \Sigma_{1i\alpha}\Sigma_{1\alpha j} + \dots \\ &= (1+\Sigma_1+\Sigma_1^2+\dots)\delta_{ij} = \frac{1}{1-\Sigma_1}\delta_{ij} \\ &\equiv B_1\delta_{ij}, \end{aligned} \tag{29}$$

$$\begin{aligned} B_{2ij} &= \delta_{ij} + \Sigma_{2ij} + \Sigma_{2i\beta}\Sigma_{2\beta j} + \dots \\ &= (1+\Sigma_2+\Sigma_2^2+\dots)\delta_{ij} = \frac{1}{1-\Sigma_2}\delta_{ij} \\ &\equiv B_2\delta_{ij}. \end{aligned} \tag{30}$$

This confirms our bonus result in subsection III D that B_{1ij} and B_{2ij} have to be diagonal. Going back to Eq. (27) and Eq. (28), we get

$$\Sigma_1 = \frac{1}{zN}\sum_{\beta_p=1}^{N} B_{2\beta_p\beta_p} = \frac{B_2}{zN}\sum_{\beta_p=1}^{N}\delta_{\beta_p\beta_p} = \frac{1}{z}B_2, \tag{31}$$

$$\Sigma_2 = \frac{1}{zN}\sum_{\alpha_p=1}^{M} B_{1\alpha_p\alpha_p} = \frac{B_1}{zN}\sum_{\alpha_p=1}^{M}\delta_{\alpha_p\alpha_p} = \frac{b}{z}B_1. \tag{32}$$

Combining the work above, we get the following equation set

$$\begin{cases} B_1 = \dfrac{1}{1-\Sigma_1} \\ B_2 = \dfrac{1}{1-\Sigma_2} \\ \Sigma_1 = \dfrac{1}{z}B_2 \\ \Sigma_2 = \dfrac{b}{z}B_1 \end{cases} \tag{33}$$

Because we are eventually interested in $B_{ij}(z) = B_{1ij}(z) = B_1(z)\delta_{ij}$, we eliminate the other three variables and get the equation of $B_1(z)$:

$$bB_1^2 - [z-(1-b)]\,B_1 + z = 0. \tag{34}$$

Solving this and plugging it into Eq. (16) and Eq. (15), we get

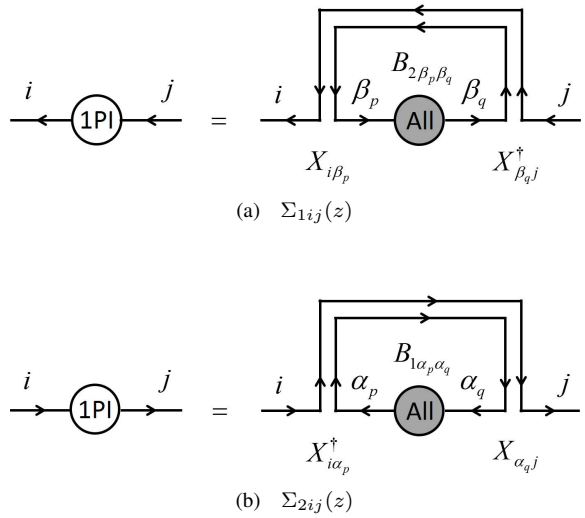


FIG. 6. 1PI diagrams (a) $\Sigma_{1ij}(z)$, and (b) $\Sigma_{2ij}(z)$, in terms of two point full diagrams

the result of $\rho(E)$

$$\begin{aligned}\rho(E) &= -\frac{1}{\pi}\lim_{\varepsilon\to 0^+}\mathrm{Im}G(E+i\varepsilon)\\ &= \frac{1}{2\pi}\frac{1}{b}\frac{\sqrt{(x_2-E)(E-x_1)}}{E}I_{\{E\in(x_1,x_2)\}}\\ &\quad +(1-\frac{1}{b})\delta(E)I_{\{b\geq 1\}},\end{aligned} \tag{35}$$

which is exactly what we want to prove (Eq. 7). Note that there are two solutions for Eq. (34), and one needs be cautious while choosing the root and taking the limit. It is a straight forward but slightly tedious procedure. To keep this paper self contained, we include this procedure as the appendix.

IV. THE CASE OF SYMMETRIC X

In this section, we prove that Marchenko-Pastur law still holds when X is symmetric complex matrix and thus not all of its elements are independent. This is important, for example, in the case of large N analysis of neutrino anarchy where the Majorana mass matrix is symmetric.

For symmetric X, we require $M = N$, thus $b = 1$. And the i.i.d. condition Eq. (1) in MP law should be regarded as for the $N(N+1)/2$ free elements only. Then the whole analysis through Section III works for symmetric X, except that we need to modify the propagator Eq.(25) into:

$$\left\langle X_{ij}X^\dagger_{kl}\right\rangle_X = \delta_{il}\delta_{jk} + \delta_{ik}\delta_{jl}. \tag{36}$$

This affects our calculation only through Eq.(27) and (28). However, by plugging (36) in, we can easily see that the new terms in these two equations vanish under $N\to\infty$

$$\begin{aligned}\Sigma_{1ij}(z) &= \sum_{\beta_p,\beta_q=1}^{N}\frac{1}{zN}(\delta_{ij}\delta_{\beta_p\beta_q}+\delta_{i\beta_q}\delta_{\beta_p j})B_{2\beta_p\beta_q}\\ &= \delta_{ij}\frac{1}{zN}\sum_{\beta_p=1}^{N}B_{2\beta_p\beta_p}+\frac{1}{zN}B_{2ji}\\ &\to \delta_{ij}\frac{1}{zN}\sum_{\beta_p=1}^{N}B_{2\beta_p\beta_p},\\ \Sigma_{2ij}(z) &= \sum_{\alpha_p,\alpha_q=1}^{M}\frac{1}{zN}(\delta_{ij}\delta_{\alpha_p\alpha_q}+\delta_{i\alpha_q}\delta_{\alpha_p j})B_{1\alpha_p\alpha_q}\\ &= \delta_{ij}\frac{1}{zN}\sum_{\alpha_p=1}^{M}B_{1\alpha_p\alpha_p}+\frac{1}{zN}B_{1ji}\\ &\to \delta_{ij}\frac{1}{zN}\sum_{\alpha_p=1}^{M}B_{1\alpha_p\alpha_p}.\end{aligned}$$

Therefore in the case of symmetric X, we will be led to the same result (Eq. 35) as in section III. And here we should take the $M = N$, i.e. $b = 1$ special case of it:

$$\rho(E) = \frac{1}{2\pi}\sqrt{\frac{4}{E}-1}\cdot I_{\{E\in(0,4)\}}. \tag{37}$$

V. CONCLUSIONS

Method with large N random matrices is greatly used in various of theoretical models. Marchenko Pastur law is a useful theorem for eigenvalue distribution of large N random matrices. We provide an alternative proof of Marchenko Pastur law using Feynman diagrams.

Appendix: Root Selection

Let us start with Eq. (34):

$$bB_1^2 - [z-(1-b)]\,B_1 + z = 0. \tag{A.1}$$

This equation gives us two analytical solutions, which for the moment, we formally write as

$$B_1(z) = \frac{z-(1-b)-r(z)}{2b}, \tag{A.2}$$

where we have used $r(z)$ to denote the square root

$$r(z) \equiv \left\{[z-(1-b)]^2 - 4bz\right\}^{\frac{1}{2}}, \tag{A.3}$$

and put in by hand a minus sign in front of it, just for future convenience. The selection of root is still undone until

we specify the branch of this multi-value function $r(z)$. The expression of $G(z)$ follows

$$G(z) = \frac{1}{M}\frac{1}{z}tr\left[B(z)\right] = \frac{1}{z}B_1(z) = \frac{1}{2b}\left\{1 - \frac{1-b}{z} - \frac{r(z)}{z}\right\}. \quad \text{(A.4)}$$

Recall that ever since Eq.(18), we have been working within the vicinity of $z = \infty$. So we need to pick the correct solution to Eq.(A.1) which is the analytically continuation of $B(z)$ from this vicinity. Checking our definition of $G(z)$ (Eq. 16) and $B(z)$ (Eq. 17), we see that both of them should be analytical at $z = \infty$, with the values

$$\lim_{z\to\infty} G(z) = 0, \quad \text{(A.5)}$$

$$\lim_{z\to\infty} B_1(z) = 1. \quad \text{(A.6)}$$

This requires $\frac{r(z)}{z}$ be analytical at $z = \infty$ with the value

$$\lim_{z\to\infty} \frac{r(z)}{z} = 1. \quad \text{(A.7)}$$

To find the form of $r(z)$ satisfying these conditions, we first notice that the two solutions to the equation $[z-(1-b)]^2 - 4bz = 0$ are both real positive due to $b > 0$. We denote them as

$$x_1 = (1-\sqrt{b})^2, \quad \text{(A.8)}$$

$$x_2 = (1+\sqrt{b})^2. \quad \text{(A.9)}$$

Then

$$[r(z)]^2 = [z-(1-b)]^2 - 4bz = (z-x_1)(z-x_2). \quad \text{(A.10)}$$

If we define

$$z - x_1 \equiv r_1 e^{i\theta_1}, \quad \text{(A.11)}$$

$$z - x_2 \equiv r_2 e^{i\theta_2}, \quad \text{(A.12)}$$

the root $r(z)$ can be written as

$$r(z) = \left\{[z-(1-b)]^2 - 4bz\right\}^{\frac{1}{2}} = \sqrt{r_1 r_2}e^{i\frac{\theta_1+\theta_2}{2}}. \quad \text{(A.13)}$$

Then specifying the branch is just to specify the values of the arguments θ_1, θ_2. For a single branch point, for example x_1, a typical assignment of θ_1 would look like Fig. 7. But any line starting from x_1 ending at ∞ can serve as the branch cut. We thus have many choices for each branch cut. However, to make $\frac{r(z)}{z}$ analytical at $z = \infty$, we have to overlap these two branch cuts, both to the left (or equivalently both to the right). The remaining freedom of globally shifting θ_1 or θ_2 by integer multiple of 2π is fixed by condition Eq. (A.7). It turns out the correct assignment (Fig. 8) is just a repetition of Fig. 7 applied to both x_1 and x_2. (If we did not put in a minus sign by hand in Eq. (A.2), Eq. (A.7) would require us to globally shift the assignment of θ_2 (or θ_1) by $\pm 2\pi$ in Fig. 8. This would result in the same minus sign for $r(z)$.)

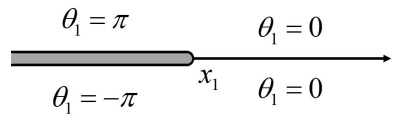


FIG. 7. Typical argument value assignment for a single branch point.

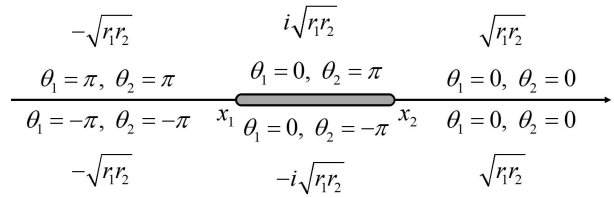


FIG. 8. Correct value assignment of arguments θ_1, θ_2 and the resulting value of $r(z)$.

Now we can compute $\rho(E)$. From the solution

$$G(z) = \frac{1}{2b}\left\{1 - \frac{1-b}{z} - \frac{r(z)}{z}\right\}, \quad \text{(A.14)}$$

and the branch structure of $r(z)$ (Fig. 8), we clearly see that $\lim_{\varepsilon\to 0^+} \mathrm{Im}G(E+i\varepsilon) = 0$ except when E falls on the branch cut of $r(z)$: $E \in (x_1, x_2)$, or E hits the pole of $G(z)$: $E = 0$. For the first case, the only contribution to $\lim_{\varepsilon\to 0^+} \mathrm{Im}G(E+i\varepsilon) = 0$ comes from $r(z)$, and from Fig. 8 we get

$$\lim_{\varepsilon\to 0^+} \mathrm{Im}G(E+i\varepsilon) \supset -\frac{1}{2b}\frac{\sqrt{r_1 r_2}}{E} = -\frac{1}{2b}\frac{\sqrt{(x_2-E)(E-x_1)}}{E} I_{\{E\in(x_1,x_2)\}}. \quad \text{(A.15)}$$

For the second case, we need to compute the residue of the pole $z = 0$

$$\mathrm{res}(G(z=0)) = \frac{1}{2b}\left\{-(1-b) - (-\sqrt{r_1 r_2})\right\} = -\frac{1}{2b}\left\{1 - b - \sqrt{(1-b)^2}\right\} = (1-\frac{1}{b})\cdot I_{\{b\geq 1\}},$$

which gives

$$\lim_{\varepsilon\to 0^+} \mathrm{Im}G(E+i\varepsilon) \supset (1-\frac{1}{b})I_{\{b\geq 1\}}\cdot\left[-\pi\delta(E)\right]. \quad \text{(A.16)}$$

Combining the two pieces, we eventually get our result

$$\rho(E) = -\frac{1}{\pi}\lim_{\varepsilon\to 0^+} \mathrm{Im}G(E+i\varepsilon) = \frac{1}{2\pi}\frac{1}{b}\frac{\sqrt{(x_2-E)(E-x_1)}}{E} I_{\{E\in(x_1,x_2)\}} + (1-\frac{1}{b})\delta(E) I_{\{b\geq 1\}}, \quad \text{(A.17)}$$

with $x_1 = (1-\sqrt{b})^2$, $x_2 = (1+\sqrt{b})^2$.

ACKNOWLEDGMENTS

This work was supported in part by the U.S. DOE under Contract DE-AC03-76SF00098, and in part by the NSF under grant PHY-1002399. The work by H.M. was also supported in part by the JSPS grant (C) 23540289, in part by the FIRST program Subaru Measurements of Images and Redshifts (SuMIRe), CSTP, Japan, and by WPI, MEXT, Japan.

[1] S. Coleman, "Aspects of symmetry," (1987).
[2] G. 't Hooft, Nucl.Phys. **B72**, 461 (1974).
[3] A. V. Manohar, , 1091 (1998), arXiv:hep-ph/9802419 [hep-ph].
[4] Y. Bai and G. Torroba, JHEP **1212**, 026 (2012), arXiv:1210.2394 [hep-ph].
[5] V. Marčenko and L. Pastur, Mathematics of the USSR-Sbornik **1**, 457 (1967).
[6] F. Götze and A. Tikhomirov, Bernoulli **10**, 503 (2004).
[7] L. J. Hall, H. Murayama, and N. Weiner, Phys.Rev.Lett. **84**, 2572 (2000), arXiv:hep-ph/9911341 [hep-ph].
[8] N. Haba and H. Murayama, Phys.Rev. **D63**, 053010 (2001), arXiv:hep-ph/0009174 [hep-ph].
[9] X. Lu and H. Murayama, (2014), arXiv:1405.0547 [hep-ph].