

概率论、费曼图以及随机矩阵理论（三）： 大N展开与随机矩阵理论

陈童

(这个小系列的三篇文章是一个整体，它试图在量子场论的理论体系之外介绍一些相关的方法和推导几个漂亮的结论。但是最后这第三篇基本上是按照A.Zee, Quantum Field Theory in a Nutshell的相关内容来讲的，本质上并没有什么新意。不过，我们这里和前两篇放在一起可以使得读者无需阅读A. Zee这本很厚的量子场论书，只需阅读本系列的前两篇小文章就能学习大N展开以及学习物理学家对随机矩阵理论的处理方法了。)

在中心极限定理的证明过程中我们看到，对于一个有 N 个随机变量的系统，它的某些问题在 N 很大的极限下会得到极大的简化。这一观察促使人们去研究各种 N 个随机变量的系统在大 N 极限下的行为，以及这些系统按照 $1/N$ 的微扰展开，这就是所谓的大 N 展开思想。在物理学中，大 N 展开思想已经有好几十年的历史发展脉络，对于今天的人来说，大 N 展开最著名的例子也许是AdS/CFT对应。不过，在这里我们并不想系统性地回顾大 N 展开的这些发展。相反，我们想介绍大 N 展开的一个重要例子，即随机矩阵理论。除了因为这一理论本身有广泛的应用以外，我们介绍它是因为，就好像中心极限定理一样，这一理论的大 N 极限是已经被了解得很清楚的，是一个漂亮的数学物理定理。

下面我们来介绍随机矩阵理论。这时候，我们要考察的随机变量是一个 $N \times N$ 的厄米矩阵 Φ ，其取值我们依然记为 ϕ ， ϕ 的第*i*行第*j*列我们记为 ϕ_j^i 。我们假定这个随机变量的概率分布如下

$$P(\phi) = \frac{1}{Z} \exp(-N \text{Tr} V(\phi)), \quad (1)$$

式中 $V(\phi)$ 是关于 ϕ 的一个实多项式，比如说 $V(\phi) = \frac{1}{2}m^2\phi^2 + g\phi^4$ 等等。

我们主要关心 ϕ 的本征值密度的期望值 $\rho(E)$ ，其定义如下

$$\rho(E) = \frac{1}{N} \langle \text{Tr} \delta(E - \Phi) \rangle = \frac{1}{N} \langle \sum_i \delta(E - E_i) \rangle, \quad (2)$$

式中 E_i 为某个按照概率(1)随机产生的 ϕ 的本征值。 $\rho(E)$ 也就是随机产生许许多多 ϕ 以后，它们的本征值的分布函数，它显然满足 $\int dE \rho(E) = 1$ 。当 N 有限时， $\rho(E)$ 很复杂很难精确地求解出来，但是，正如我们将要看到的，在 $N \rightarrow +\infty$ 的大 N 极限下， $\rho(E)$ 将趋向于一个简单的光滑函数。我们下面要做的事情就是找到这个光滑函数，它就类似于我们在中心极限定理中的高斯分布。

注意到 $\lim_{\epsilon \rightarrow 0} \text{Im}(\frac{1}{x+i\epsilon}) = -\pi\delta(x)$ ，我们可以引入一个更容易计算的函数 $G(z)$ ，其定义为

$$G(z) = \frac{1}{N} \langle \text{Tr} \frac{1}{z - \Phi} \rangle. \quad (3)$$

很显然， $\rho(E) = -(1/\pi) \lim_{\epsilon \rightarrow 0} \text{Im} G(E + i\epsilon)$ 。另外，注意到我们的概率分布(1)在幺正变换下是不变的，即 $P(\phi) = P(U^\dagger \phi U)$ ，这里 U 为一个任意的 $N \times N$ 的幺正矩阵。因此我们必然有

$$G_j^i(z) = \left\langle \left(\frac{1}{z - \Phi} \right)_j^i \right\rangle = \delta_j^i G(z). \quad (4)$$

1 魏格纳半圆定律

下面我们考虑一种简单的情况，即 $V(\phi) = \frac{1}{2}m^2\phi^2$ 的情况。这时候由高斯积分我们容易有

$$\langle \Phi_j^i \Phi_l^k \rangle = \frac{1}{Z} \int [d\phi] \phi_j^i \phi_l^k \exp \left(-N \frac{1}{2} m^2 \sum_{p,q} \phi_q^p \phi_p^q \right) = \delta_l^i \delta_j^k \frac{1}{Nm^2}. \quad (5)$$

这就是高斯分布情况下随机矩阵理论中两个外点进行配对所需满足的规则。由于 Φ_j^i 带有上下两个不同的指标，因此我们常常用如图(1)所示的双线来表示这种配对。注意图中箭头的方向总是从下指标指向上指标的，并请注意图中两根线的箭头方向相反。

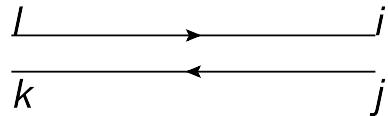


Figure 1: The pairing between Φ_j^i and Φ_l^k with contribution $\delta_l^i \delta_j^k \frac{1}{Nm^2}$. The i, j end of the double lines stands for Φ_j^i , the k, l end stands for Φ_l^k .

现在，我们可以计算 $G_j^i(z)$ 了，注意到奇数个 Φ 的期望值一定为零，我们可以将 $G_j^i(z)$ 按照 $1/z$ 泰勒展开为

$$G_j^i(z) = \sum_{n=0}^{\infty} \frac{1}{z^{2n+1}} \langle (\Phi^{2n})_j^i \rangle \quad (6)$$

$$= \delta_j^i \frac{1}{z} + \frac{1}{z^3} \langle (\Phi^2)_j^i \rangle + \frac{1}{z^5} \langle (\Phi^4)_j^i \rangle + \dots \quad (7)$$

首先让我们来看这个展开中的第二项，即 $\frac{1}{z^3}\langle\Phi_k^i\Phi_j^k\rangle$ （当一个表达式中的一个上指标和一个下指标用同样的字母表示时，我们默认对这个指标进行求和，也称之为将这两个指标进行收缩），稍微想一下就能发现，这一项可以用如下的图形(2)来表示。

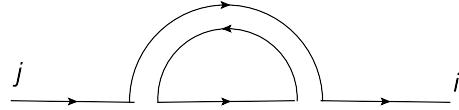


Figure 2: The diagram for $\frac{1}{z^3}\langle\Phi_k^i\Phi_j^k\rangle$.

图中每一根单线的贡献是 $1/z$ ，当用一根单线将双线的两端连起来的时候表示要将这两端的两个指标进行收缩。因此，这幅图的总贡献为 $\delta_j^i N \frac{1}{Nm^2}$ ，总的来说是在 N 的零次方阶，我们记为 $O(N^0)$ 。注意这里额外的那个 N ，它显然来自于上下指标的收缩，也就是对应于图(2)中那个闭合的回路。稍微想一下就能发现，任何闭合回路总意味着一个这样的收缩，因此会额外贡献一个因子 N 。

下面我们来考察展开式(7)中的第三项，也就是 $\frac{1}{z^5}\langle\Phi_k^i\Phi_l^k\Phi_h^l\Phi_j^h\rangle$ ，这涉及到四个矩阵，根据高斯积分两两配对的规则，它有多种配对的可能性。暂时不考虑这四个 Φ 如何配对，仅仅注意到这四个 Φ 的上下指标收缩关系，我们容易知道相应的图形必然具有如下结构，如图(3)。图中的阴影部分表示双线间的可能配对。

现在，让我们来考虑双线间的可能配对形式。它给出如图(4)所示的3种可能性。

图(4)中的每一幅图都有两根双线，每根双线贡献一个 $1/N$ 。但是图(4)中的(a)、(c)两图和(b)图有根本性的不同，(a)、(c)两图包含

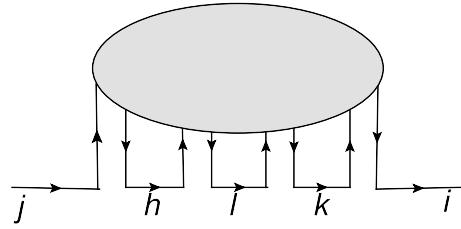


Figure 3: The diagram for $\frac{1}{z^5} \langle \Phi_k^i \Phi_l^k \Phi_h^l \Phi_j^h \rangle$.

了两个闭合回路，每个闭合回路会额外贡献一个 N ，所以总的来说(a)、(c)两图是 $O(N^0)$ 阶的。但是(b)图完全可以一笔画出，它不包含任何闭合回路，所以总的来说，(b)图是 $O(1/N^2)$ 阶的。也就是说，在大 N 极限下，(b)图的贡献完全可以忽略不计。实际上，(b)图通常被称为非平面图，也就是说当它在二维平面上画出来的时候必然要涉及一些交叉线。相反，(a)、(c)这样的图被称之为平面图，它们可以直接在二维平面上画出来而不涉及到任何交叉线。图(4)其实示例了一条一般规则，即任何非平面图相对于平面图来说总有更少的回路，由于每一个闭合回路都会额外贡献一个 N ，因此非平面图在大 N 展开的时候就代表 $1/N$ 的更高阶贡献。特别的，在 $N \rightarrow \infty$ 的大 N 极限下，我们可以忽略掉所有的非平面图。正因为如此，大 N 极限有时候也被称之为平面图极限。

下面我们考虑完整的 $G_j^i(z)$ 在平面图极限下的结构，这就要考虑泰勒展开(7)的所有阶。为此让我们首先观察一下图(4)的(a)、(c)两图的特征，很容易发现，(a)图是将同一个组成单元重复两次而成，而(c)图是在图(2)的基础上嵌套一根双线而成。稍微尝试一下就能

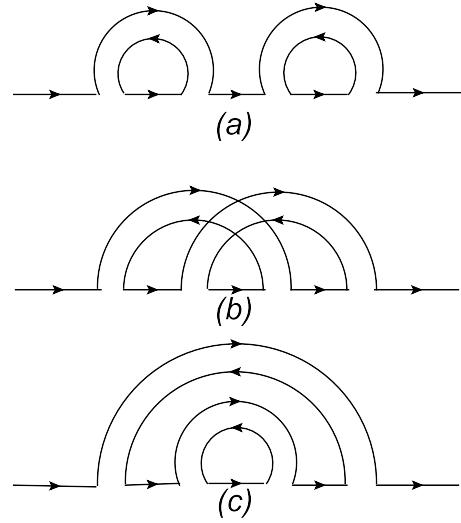


Figure 4: The diagrams for $\frac{1}{z^5} \langle \Phi_k^i \Phi_l^k \Phi_h^l \Phi_j^h \rangle$.

发现，在平面图极限下，这两个特征是泰勒展开所有阶共有的，也即是说，在平面图极限下，任何对 $G_j^i(z)$ 有贡献的图都可以通过重复和嵌套这两种方式来生成。假定通过嵌套生成的所有图的贡献之和为 $S(z)$ ，并将两种操作生成的所有图的贡献之和记为 $G(z)$ (因为 $G_j^i(z) = \delta_j^i G(z)$)，如下图(5)所示。

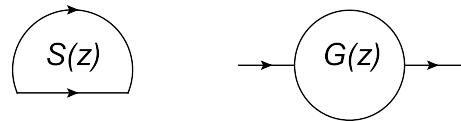


Figure 5: $S(z)$ and $G(z)$.

很清楚， $G(z)$ 是将 $S(z)$ 不断重复而成，反过来， $S(z)$ 是在 $G(z)$ 的基础上嵌套一根双线而成，用图形来表示就是(6)。从这幅图中我们很容易读出来

$$\begin{aligned} G(z) &= \frac{1}{z} + \frac{1}{z}S(z)\frac{1}{z} + \frac{1}{z}S(z)\frac{1}{z}S(z)\frac{1}{z} + \dots \\ &= \frac{1}{z - S(z)}. \end{aligned} \quad (8)$$

另外，从图中也可以看出来，从 $G(z)$ 通过嵌套得到 $S(z)$ 的过程我们加上了一根双线，其贡献为 $1/(Nm^2)$ ，但同时也增加了一个回路，其贡献为 N ，因此我们有

$$S(z) = \frac{1}{m^2}G(z). \quad (9)$$

通过(8)和(9)这两个方程，我们容易解得

$$G(z) = \frac{m^2}{2}\left(z - \sqrt{z^2 - \frac{4}{m^2}}\right). \quad (10)$$

式中在二次方程两个根的±号中，我们选择了-号，这是因为只有这样相应的 $G(z)$ 才能满足在 $z \rightarrow \infty$ 时 $G(z) \rightarrow 1/z$ 。

由 $G(z)$ 的表达式(10)，我们很容易得到本征值分布 $\rho(E)$ ，

$$\rho(E) = \frac{m^2}{2\pi} \sqrt{\frac{4}{m^2} - E^2}, \quad (11)$$

式中 E 的取值在区间 $[-2/m, +2/m]$ 之内。这就是魏格纳半圆定律。

2 戴森气体

魏格纳半圆定律是大 N 极限下随机矩阵本征值分布的一个简单结果。但是这个结果只适用于 $V(\phi) = \frac{1}{2}m^2\phi^2$ 的情形。这一节我们就

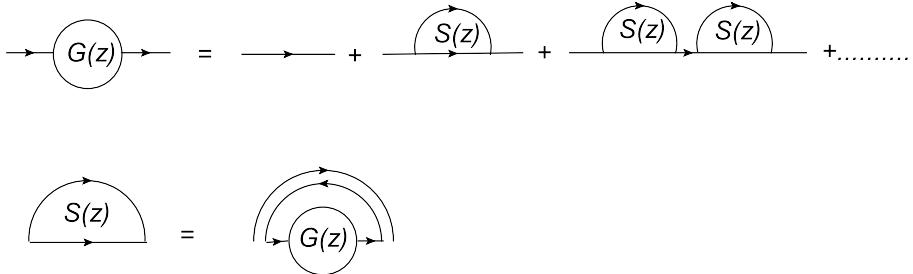


Figure 6: The diagrams for $G(z)$ and $S(z)$.

是要把类似的结论推广到 $V(\phi)$ 是一个一般性的多项式的情形，比方说 $V(\phi) = \frac{1}{2}m^2\phi^2 + g\phi^4$ 这样的情形。如果用按照 g 进行微扰展开的方式来处理问题，我们就能发现，这时候，顶角 $g\phi^4$ 的引入使得我们需要考虑的典型平面图为如(7)所示的样子。很清楚，我们很难通过直接将所有类似这样的图的贡献都加起来以得到一个 $G(z)$ 的表达式，所以我们可能需要用图形之外的方法来处理现在的一般性情形。

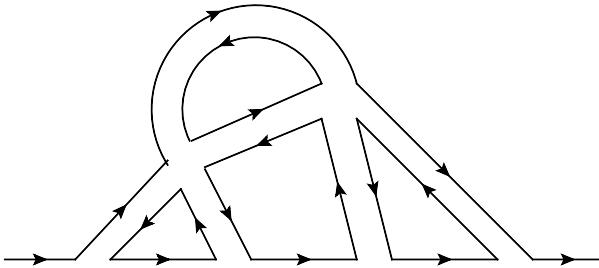


Figure 7: A diagram with two vertexes.

为此让我们来考察一个观测量 $\text{Tr}(\mathcal{O}(\phi))$ 的期望值，这里 $\mathcal{O}(\phi)$ 表

示 ϕ 的一个函数。很清楚， $\text{Tr}(\mathcal{O}(\phi))$ 在幺正变换 $\phi \rightarrow U^\dagger \phi U$ 下是不变的。这个期望值由下式计算

$$\langle \text{Tr}(\mathcal{O}(\phi)) \rangle = \frac{1}{\mathcal{Z}} \int [d\phi] \text{Tr}(\mathcal{O}(\phi)) e^{-N \text{Tr} V(\phi)}. \quad (12)$$

前面我们已经说过了概率分布 $P(\phi)$ 在幺正变换下是不变的， $\text{Tr}(\mathcal{O}(\phi))$ 也是幺正不变的，值得讨论的是积分测度 $[d\phi]$ ，不过可以证明如果将它自然地定义成 $[d\phi] = \prod_{i,j} d\phi_j^i$ ，则它在幺正变换下也是不变的。

另一方面，对于任何厄米矩阵 ϕ ，如果将其本征值所构成的对角矩阵记为 $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ ，则 ϕ 必定可以分解成 $\phi = U^\dagger \Lambda U$ 的形式，式中 U 为某个幺正矩阵。那么积分测度 $[d\phi]$ 必定可以重写成 $[d\phi] = J \prod_i d\lambda_i [dU]$ ，式中 J 为雅可比行列式， $[dU]$ 表示对幺正矩阵 U 的积分测度。由于积分表达式(12)的被积函数是幺正变换不变的，因此对 $[dU]$ 的积分是平凡，不妨将之归一化为1，如此一来，积分(12)就可以重写成，

$$\langle \text{Tr}(\mathcal{O}(\phi)) \rangle = \frac{1}{\mathcal{Z}} \int J \prod_i d\lambda_i \sum_j (\mathcal{O}(\lambda_j)) e^{-N \sum_k V(\lambda_k)}. \quad (13)$$

这个表达式中唯一还不清楚的就是那个雅可比矩阵 J 。

为了求出雅可比矩阵 J ，我们不妨考虑 ϕ 在对角矩阵邻域内的情形，即 $\phi = e^{-i\varepsilon} \Lambda e^{i\varepsilon}$ ，其中 ε 为一个无穷小厄米矩阵，因此有 $\phi = \Lambda - i[\varepsilon, \Lambda]$ ，由这个式子可以看出，对 Λ 的积分(即 $[d\Lambda] = \prod_i d\lambda_i$)带来的雅可比行列式为1，同时由于矩阵元 $[\varepsilon, \Lambda]_{ij} = (\lambda_i - \lambda_j)\varepsilon_{ij}$ ，它带来的积分测度为 $\prod_{i,j} [(\lambda_i - \lambda_j)d\varepsilon_{ij}] = \prod_{i < j} (\lambda_i - \lambda_j)^2 [d\varepsilon]$ ，其中积分测度 $[d\varepsilon] = \prod_{i,j} d\varepsilon_{ij}$ 就是测度 $[dU]$ 在恒等矩阵邻域内的具体形式。因此，我们有，雅可比行列式 $J = \prod_{i < j} (\lambda_i - \lambda_j)^2$ 。将之代入积分

式(13)我们就有

$$\langle \text{Tr}(\mathcal{O}(\phi)) \rangle = \frac{1}{\mathcal{Z}} \int \prod_i d\lambda_i \sum_j (\mathcal{O}(\lambda_j)) \exp [-NS(\lambda_1, \lambda_2, \dots, \lambda_N)]. \quad (14)$$

式中 $S(\lambda_1, \lambda_2, \dots, \lambda_N) = \sum_i V(\lambda_i) - \frac{1}{N} \frac{1}{2} \sum_{i \neq j} \log(\lambda_i - \lambda_j)^2$ 。可以看到，现在问题变成了一个 N 粒子系统的问题，粒子的坐标就是 $\lambda_i, i = 1, 2, \dots, N$ 。这 N 个粒子处在一个势阱 $V(\lambda)$ 中，并且任意两个粒子之间有排斥势 $-\frac{1}{N} \log(\lambda_i - \lambda_j)^2$ 的相互作用。这 N 个粒子就构成某种有相互作用的气体，称之为戴森气体。

从表达式(14)很容易看出，在大 N 极限下，只有函数 $S(\lambda_1, \lambda_2, \dots, \lambda_N)$ 的极小值对积分有贡献，因此有 $\frac{\partial}{\partial \lambda_i} S(\lambda_1, \lambda_2, \dots, \lambda_N) = 0, i = 1, 2, \dots, N$ ，由此就得到如下方程

$$V'(\lambda_i) = \frac{2}{N} \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j}. \quad (15)$$

当 $N \rightarrow \infty$ 时，这些本征值就会趋向于一个连续谱，引入谱密度 $\rho(\lambda)$ （它就是我们前面引入的本征值分布函数 $\rho(E)$ 的大 N 极限，因此满足 $\int d\lambda \rho(\lambda) = 1$ ），我们就可以把上面这个方程重写为

$$V'(\lambda) = 2\mathcal{P} \int d\mu \frac{\rho(\mu)}{\lambda - \mu}. \quad (16)$$

式中 \mathcal{P} 表示取积分主值。

从方程(16)容易看出，大 N 极限下的本征值谱必然分布在一个有限的区间之内，因为在允许 $|\lambda| \rightarrow \infty$ 的情况下，方程(16)右边趋于 $\sim 1/\lambda$ ，因此是无法和左边相等的。下面我们将把这个本征值的分布区间记为 A 。

下面我们引入解析函数 $G(z) = \int_A d\mu_{z-\mu}^{\rho(\mu)}$, 其中变量 z 定义在整个复平面上。很容易看出来 $\rho(\lambda) = -\frac{1}{\pi} \text{Im}G(\lambda + i\epsilon)$, 因此这里的 $G(z)$ 实际上就是我们前面引入的 $G(z)$ 函数的大 N 极限, $G(z)$ 现在的这种表达形式称之为谱表示, 从这种表达形式可以看出来, $G(z)$ 除了在区间 A 上有一条割线以外, 它在整个复平面都是解析的。很显然方程(16)告诉我们的不过是 $\text{Re}G(\lambda + i\epsilon) = \frac{1}{2}V'(\lambda)$ (在区间 A 上)。由这个条件, 加上 $G(z)$ 的解析性, 再加上 $z \rightarrow \infty$ 时 $G(z) \rightarrow 1/z$ 的条件, 我们就能唯一性地确定解析函数 $G(z)$, 由此就能求出 $\rho(\lambda)$ 。

比方说, 假如 $V(\lambda) = \frac{1}{2}m^2\lambda^2 + g\lambda^4$, 这时候由于 $V(\lambda)$ 是一个偶函数, 我们可以知道, 谱分布 $\rho(\lambda)$ 也必然是一个偶函数, 因此区间 A 必定可以取成 $[-a, a]$ 的形式, 其中 a 是待定的。魏格纳半圆定律启发我们可以试探性地将 $G(z)$ 取成如下形式

$$G(z) = \frac{1}{2}V'(z) - P(z)\sqrt{z^2 - a^2}, \quad (17)$$

其中 $P(z)$ 是一个待确定的多项式。很显然, 这时候

$$\rho(\lambda) = \frac{1}{\pi}P(\lambda)\sqrt{a^2 - \lambda^2}, \quad (18)$$

由于 $\rho(\lambda)$ 是一个偶函数, 所以 $P(z)$ 必定是一个偶多项式。又由于 $z \rightarrow \infty$ 时 $G(z) \rightarrow 1/z$, 可知 $P(z)$ 是一个2次偶多项式, 因此包含两个待定系数。再加上前面待定的 a , 我们就有3个待定参数。它们可以这样来确定, 由于 $z \rightarrow \infty$ 时 $G(z) \rightarrow 1/z$, 所以在 $z \rightarrow \infty$ 时 $G(z)$ 渐进式的 z^3, z^1 项系数都要消去, 而 $1/z$ 项的系数必定要等于1, 这就给出了三个方程, 刚好可以确定3个待定参数。这就完成了我们想要的求解, 相应的(18)式就是魏格纳半圆定律的推广。

得到 $\rho(\lambda)$ 后，观察量 $\text{Tr}\mathcal{O}(\phi)$ 的期望值在大 N 极限下就可以写成

$$\frac{1}{N} \langle \text{Tr}(\mathcal{O}(\phi)) \rangle = \frac{1}{N} \sum_i \mathcal{O}(\lambda_i) = \int d\lambda \rho(\lambda) \mathcal{O}(\lambda). \quad (19)$$