

Введение в случайные матрицы

Александр Артемьев

ОП «КТП, теория струн и мат. физика», курс «Избранные главы теоретической и математической физики»

1. Что и зачем?

Теория случайных матриц (или *матричные модели*, это одно и то же) ставит своим вопросом вычисление средних значений функций от элементов матрицы из некоторого класса, которые распределены по какому-то вероятностному закону. Для матриц маленького размера (по крайней мере принципиально) любое среднее можно вычислить в лоб; обычно людей интересует, однако, случай, когда размер матрицы $N \gg 1$. Среди физических приложений матричных моделей можно выделить следующие:

- 1) во многих физических системах со сложными гамильтонианами и большим числом элементарных степеней свободы (например, в ядерной физике) плотность уровней энергии может быть приближена распределением собственных значений большой случайной матрицы. В этом контексте матричные модели впервые появились в физике;
- 2) матричные модели — игрушечная модель КТП (в частности, калибровочных теорий): они предоставляют возможность изучить и лучше понять многие методы квантовой теории поля (такие, как $1/N$ разложение, петлевые уравнения Макеенко-Мигдала, обычная теория возмущений, тождества Уорда...) в простом случае;
- 3) наконец, матричные модели дают точное описание некоторых моделей КТП: некоторые наблюдаемые в нетривиальных квантовых теориях поля (таких, как суперсимметричный Янг-Миллс или двумерная квантовая гравитация) допускают вычисление с помощью матричных моделей с большим N .

Для нас больше всего будет актуален второй пункт в этом списке.

2. Эрмитова матричная модель. Теория возмущений, $1/N$ разложение.

Первое, что нужно выбрать для формулировки матричной модели — класс матриц, в котором мы работаем («ансамбль»). Мы из соображений простоты будем рассматривать *унитарный ансамбль* — так называется матричная модель, где мы усредняем по *эрмитовым* матрицам $H^\dagger = H$ (терминология слегка запутывает, но она стандартная). В статистических приложениях выбор ансамбля определяется симметриями задачи; другие часто используемые ансамбли — ортогональный (из симметричных вещественных матриц) и симплектический (из антисимметричных).

Усреднение по всем эрмитовым матрицам можно представить в виде интегрирования по всем матричным элементам; они задаются $N^2 - N$ вещественными числами H_{ii} , $\text{Re } H_{ij}$, $\text{Im } H_{ij}$ ($i < j$) (условие эрмитовости даёт, что H_{ii} вещественно, $\text{Re } H_{ij} = \text{Re } H_{ji}$, $\text{Im } H_{ij} = -\text{Im } H_{ji}$). Общая формула для вычисления средних

$$\langle f(H) \rangle = \int [DH] \rho(H) f(H) \equiv \prod_{i=1}^N \left(\int dH_{ii} \right) \prod_{j < i} \left(\int d\text{Re } H_{ij} d\text{Im } H_{ij} \right) \rho(H) f(H)$$

Достаточно общим случаем является выбор плотности вероятности в виде экспоненты от «single-trace» функции: $p(H) = Z^{-1} \exp(-N \text{tr } V(H))$, где Z — статсумма (нормировочный фактор, выбранный так, чтобы $\langle 1 \rangle = 1$), а V — произвольная функция, называемая потенциалом (можно думать

про неё как про полином: $V(H) = \frac{g_2 H^2}{2!} + \frac{g_3 H^3}{3!} + \dots$). Легко написать ответы для многих средних в случае квадратичного потенциала, когда наш матричный интеграл просто многомерный гауссовый. Так, самое простое среднее

$$\langle H_{ij} H_{kl} \rangle = \frac{1}{N g_2} \delta_{il} \delta_{kj}$$



Упражнение. Проверьте эту формулу. Как она отличается от аналогичного ответа в ортогональном ансамбле?

Средние от более высоких степеней $\langle H_{i_1 j_1} \dots H_{i_n j_n} \rangle$ можно вычислить по *теореме Вика*. Согласно этой теореме они равны сумме всех $\frac{(2n)!}{2^n n!}$ возможных «свёрток» — произведений попарных средних. Если наша модель негауссова, но константы g_n , $n \neq 2$ маленькие, можно раскладывать соответствующие слагаемые в экспоненте в ряд по этим константам и искать различные средние порядок за порядком по g_n , вычисляя при помощи теоремы Вика каждое слагаемое. Это — обычная теория возмущений.

Каждое слагаемое в таком разложении удобно представлять в виде картинок — «диаграмм с двойными линиями». Каждому H в корреляционной функции соответствуют две линии — по одной на каждый индекс i, j ; они соединяются, если по соответствующему индексу идёт матричное умножение (т.е. соответствующие H свёрнуты друг с другом). Каждой замкнутой линии соответствует дополнительный фактор N , а каждому «пропагатору» — $1/N$ согласно формуле для двухточечного среднего. Примеры диаграмм для матричной модели, где не равны нулю только g_2 и g_4 , появляющихся в вычислении $\langle \text{tr } H^2 \rangle$, приведён на рисунке 1.

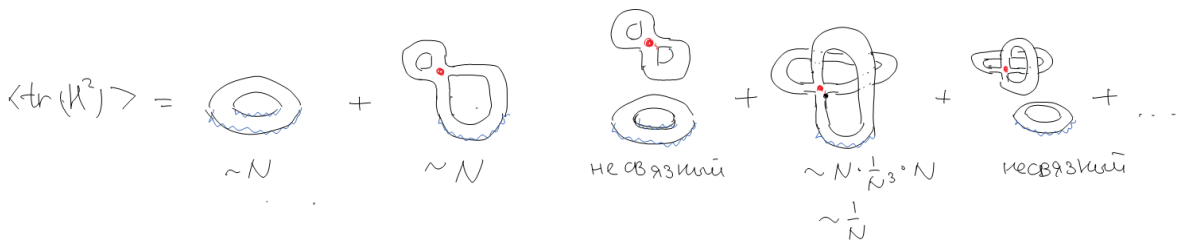


Рис. 1. Диаграммы для $\text{tr } H^2$

Удобно классифицировать диаграммы, чтобы не считать лишнего. Во-первых, есть несвязные диаграммы (состоящие из нескольких кусков); они всегда выглядят как произведение (связная, дающая вклад в $\text{tr } H^2$) \times (дающая вклад в статсумму — нульточечное среднее). Можно показать, что учёт всех таких диаграмм по сравнению с вычислением только вклада от связных просто приведет к сокращению знаменателя Z (который вообще тоже надо было бы отдельно вычислять по теории возмущений).



Упражнение. Убедитесь в этом: покажите, что полный диаграммный ряд для коррелятора можно представить как (связные диаграммы) \times (статсумма теории с негауссовым потенциалом)

Среди связных диаграмм есть диаграммы типа 1, 2 и диаграммы типа 4. Результат их вычисления показывает, что в пределе больших N диаграмма 4 подавлена по сравнению с 1, 2 фактором $1/N^2$. На уровне картинок они отличаются тем, что первые две можно нарисовать без самопересечений на плоскости, а последнюю нет. Зато её можно нарисовать на торе.



Упражнение. Эйлерова характеристика графа определяется как $\chi = V - E + F$, где V — число вершин, E — ребер а F — граней соответственно. Рассмотрим некоторую диаграмму для $\text{tr } H^k$, которая при превращении всех двойных линий в одинарные переходит в граф с эйлеровой характеристикой χ . Выразите степень N , которой пропорционально выражение для такой диаграммы, через k и χ .

Получаем, что разложение по топологиям совпадает с $1/N$ разложением: чем больше «ручек» на поверхности, на которой «живёт» диаграмма, тем сильнее подавлен её вклад в среднее при $N \rightarrow \infty$. Чтобы получить главное приближение к ответу при $N \gg 1$, можно просуммировать все планарные диаграммы (без самопересечений); многие ответы в этом лидирующем порядке можно найти аналитически для потенциала общего вида! Таким образом, $1/N$ разложение является альтернативной схемой приближений, в которой ответ в фиксированном порядке может быть непertурбативным по константам g_n .

Такое разложение по топологиям наводит на мысли об интерпретации матричных модели как дискретной двумерной квантовой гравитации; в некотором смысле проведения «суммирования по всем случайным двумерным поверхностям». Мы не будем дальше развивать эту идею.

3. «Вигнеровский полукруг» для гауссовой модели с помощью диаграммной техники.

Вернёмся обратно к гауссовой модели (и будем везде класть $g_2 = 1$ для простоты): несмотря на силу теоремы Вика, все равно есть интересные средние, вычисление которых нетривиально и в ней. Одной из них является (нормированная) плотность распределения собственных значений (в главном порядке по $1/N$ его на самом деле достаточно, чтобы найти среднее от любой их функции). Обозначим $\lambda_i(H)$, $i = 1, \dots, N$ — собственные числа матрицы H , тогда среднее, которое мы ищем, можно записать формулой

$$\rho(E) = \frac{1}{N} \sum_{i=1}^N \langle \delta(E - \lambda_i) \rangle = \frac{1}{N} \langle \text{tr} \delta(E \times \mathbb{I}_{N \times N} - H) \rangle$$

где $\delta(x)$ — дельта-функция Дирака (хотя и производится суммирование сингулярных выражений, после усреднения под интегралом ρ получается хорошей гладкой функцией; единичную матрицу перед E далее мы будем опускать). Такая плотность удовлетворяет $\int \rho(E) dE = 1$. Альтернативно (и более удобно для пертурбативного вычисления) можно воспользоваться формулой Сохоцкого

$$\frac{1}{x + i0} = \mathcal{P} \frac{1}{x} - i\pi\delta(x)$$

и искать среднее $\text{Im} \langle \text{tr} \frac{1}{E - H + i0} \rangle$. Раскладывая знаменатель по H в ряд, мы находим

$$R_{ij} \equiv \left\langle \left(\frac{1}{E + i0 - H} \right)_{ij} \right\rangle = G_{ij} + \langle G_{ik} H_{kl} G_{lj} \rangle + \langle (GHGHG)_{ij} \rangle + \dots, \quad G_{ij} \equiv \delta_{ij} \frac{1}{E + i0}$$

Проводя усреднение, мы снова можем рисовать соответствующие диаграммы (например, см. рисунок 2: фиолетовыми линиями обозначены G_{ij}).



Рис. 2. Диаграммы для R_{ij}

Формульно этот рисунок записывается как

$$R_{ij} = G_{ij} + \frac{1}{N} \text{tr} G(G^2)_{ij} + \left(\frac{1}{N^2} (\text{tr} G)^2 (G^3)_{ij} + \frac{1}{N^2} (\text{tr} G)(\text{tr} G^2)(G^2)_{ij} + \frac{1}{N^2} (G^5)_{ij} \right) + \dots$$

Во всех слагаемых, кроме последнего, наличие следа убьёт факторы $1/N$; последнее, так как в нём имеются пересекающиеся линии, подавлено. Если взглядеться в картинку, можно понять, что диаграммы без пересечений в каждом следующем порядке (по количеству нужных свёрток) даются либо последовательно нарисованными диаграммами предыдущих порядков (как диаграмма 3 на рис. 2), либо диаграммой предыдущего порядка, «накрытой» ещё одной «аркой» (как диаграммы



Рис. 3. «Уравнение Швингера-Дайсона»

2 или 4). Такая простая структура отсутствует, если мы разрешаем самопересечения! Но она позволяет в лидирующем порядке записать следующее равенство: в картинках её можно выразить рисунком 3.

В формулах

$$R_{ij} = G_{ij} + G_{ik} \Sigma_{kl} G_{lj} + (G \Sigma G \Sigma G)_{ij} + \dots = (G^{-1} - \Sigma)_{ij}^{-1}$$

$$\Sigma_{ij} = \frac{1}{N} \text{tr} R \delta_{ij}$$

Аналогичные равенства в теории поля, помогающие суммировать подклассы диаграмм как геометрическую прогрессию, называют уравнениями Швингера-Дайсона. В нашем случае такое пересуммирование даёт решаемое (квадратное) уравнение на R : в правой части стоит матрица, пропорциональная единичной, значит, $R_{ij} = r \delta_{ij}$, а из самого уравнения следует, что она удовлетворяет

$$r(E - r) = 1 \Rightarrow r = \frac{E}{2} - \sqrt{\frac{E^2}{4} - 1}$$

где нужный корень выбран, исходя из правильного поведения при больших E . Такое решение имеет ненулевую мнимую часть при $|E| < 2$; её наличие говорит нам о том, что на этом отрезке плотность ρ ненулевая и равна

$$\rho(E) = \frac{1}{\pi} \sqrt{1 - \frac{E^2}{4}}$$

В качестве нетривиальной проверки нашего вычисления заметим, что автоматически $\int \rho(E) dE = 1$, как и должно быть. График этой функции имеет форму полукруга (точнее, полуэллипса); такой результат для предельной формы распределения плотности собств. значений впервые был получен Вигнером и называется Вигнеровским полукругом.

Интересным артефактом этого приближения является то, что $\rho(E)$ сконцентрирована на конечном отрезке и строго равна нулю вне его. Это, конечно, не может быть верно ни для какого конечного N : хотя вероятность наблюдения больших собственных чисел (экспоненциально) мала, она ненулевая. Это может приводить к интересным эффектам («фазовым переходам») даже в пределе $N \rightarrow \infty$ для «достаточно плохо» себя ведущих усредняемых выражений (например, экспоненциально больших типа $\exp(\beta H)$).



Упражнение*. Разложение ответа для r в ряд по E даёт нам ряд из лидирующих по $1/N$ вкладов в $\text{tr} H^n$. Попробуйте получить их комбинаторно, посчитав все диаграммы без самопересечений в заданном порядке.

4. Переход к интегрированию по собственным значениям. Уравнение на плотность собственных значений для произвольного потенциала.

Рассмотрим альтернативный метод для вычисления средних от функций f , которые зависят только от собственных значений λ . Сначала заметим некоторые свойства меры $[DH]$ в нашем матричном интеграле. Про неё удобно думать в терминах метрики на пространстве эрмитовых матриц: вообще, если на некотором пространстве с координатами x^i задана метрика

$$ds^2 = g_{ij}(x_k) dx^i dx^j$$

то на этом пространстве появляется мера (форма объема) согласно $dV = \sqrt{\det g(x)} \prod dx_i$. Различные симметрии меры (возможные замены координат с единичным якобианом перехода) можно

искать как симметрии метрики: например, если g_{ij} константы, и мера и метрика инвариантны по отношению к сдвигам $x^i \rightarrow x^i + \epsilon^i$. Для эрмитовых матриц метрика, дающая меру $[DH]$, записывается просто как

$$ds_H^2 \equiv \langle dH, dH \rangle = \text{tr} (dH dH)$$

Можно проверить, что она инвариантна при замене переменных $H \rightarrow U^\dagger H U$, U — унитарная матрица (вообще это достаточно легко проверить и для меры непосредственно).

Далее, из линейной алгебры известно, что любую эрмитову матрицу можно параметризовать как $H = U^{-1} \Lambda U$, где U — унитарная ($U^\dagger = U^{-1}$), а $\Lambda = \text{diag} (\lambda_1, \dots, \lambda_n)$. Хотя явные выражения параметров, задающих U и λ , через элементы матрицы весьма громоздки, хотелось бы перейти к интегрированию по ним от той меры, которую мы имеем. В силу инвариантности меры $[DH]$, замеченной ранее, после такой замены переменных мы должны получить что-то вроде

$$\int [DH] = \int [DU] \left(\prod_{i=1}^n \int d\lambda_i \right) J(\lambda_k)$$

где $[DU]$ — мера на пространстве унитарных матриц, которая инвариантна при замене $U \rightarrow UX$ или $U \rightarrow XU$, X унитарная. Мера с такими свойствами известна как *мера Хаара* на группе. Явное выражение для неё весьма громоздко; зато соответствующая инвариантная метрика пишется легко:

$$ds_U^2 \equiv \langle U^{-1} dU, U^{-1} dU \rangle = \text{tr} (dU U^{-1} dU U^{-1})$$

Точный вид $[DU]$ нас не интересует, потому что от неё мы избавимся, поделив на статсумму. А вот якобиан $J(\lambda_k)$ актуален.



Упражнение*. Для вычисления этого якобиана можно подставить вариацию H через вариации U и Λ в метрику ds_H^2 и выразить её как

$$ds_H^2 = \text{tr} (d\Lambda, d\Lambda) + (U^{-1} dU)_{ij} F_{ij,kl} (U^{-1} dU)_{kl}$$

Тогда в некотором правильном смысле якобиан будет детерминантом оператора F . Это вычисление слегка громоздко, но реализуемо; попробуйте это сделать.

Проще воспользоваться т.н. *трюком Фаддеева-Попова*. Рассмотрим следующий матричный интеграл:

$$\Delta(H)^{-1} = \int [DU] \prod_{i < j} \delta^{(2)} ([U^\dagger H U]_{ij}), \quad \delta^{(2)}(x) \equiv \delta(\text{Im } x) \delta(\text{Re } x)$$

В силу инвариантности меры $[DU]$, указанной ранее, ясно, что $\Delta(H)$ зависит только от собственных значений λ . Поэтому для его вычисления будем считать, что $H = \text{diag} (\lambda_1, \dots, \lambda_N)$. Вычисление проводится с помощью формулы

$$\delta(f(x)) = \frac{\delta(x - x_0)}{|f'(x_0)|}, \quad x_0 : f(x_0) = 0$$

которая говорит, что для вычисления интеграла от «сложной» дельта-функции достаточно знать её поведение вблизи точки, где обнуляется её аргумент, в первом порядке по отклонению. В нашем случае для диагонального H аргументы дельта-функций обнуляются при $U = 1$. Вблизи этой точки ($U = 1 + a_{ij}$, $a_{ij} \ll 1$ и антиэрмитовая) мера Хаара

$$[DU] \approx \prod_{i < j} d\text{Re } a_{ij} d\text{Im } a_{ij}$$

а аргументы дельта-функций равны $a_{ij}(\lambda_i - \lambda_j)$. Таким образом,

$$\Delta(H)^{-1} = \prod_{i < j} (\lambda_i - \lambda_j)^{-2}$$

Корень из этого выражения известен как определитель Вандермонда. Для чего же мы провели это вычисление? Затем, чтобы теперь вставить под наш матричный интеграл «умную единицу» и провести следующую последовательность выкладок:

$$\begin{aligned} \int [DH] \rho(\lambda) f(\lambda) &= \int [DH] \Delta(H) \int [DU] \prod_{i < j} \delta^{(2)}([U^\dagger H U]_{ij}) = (\phi = U^\dagger H U) \\ &= \left(\prod_{i < j} \int d^2 \phi_{ij} \delta^{(2)}(\phi_{ij}) \right) \times \left(\prod_{i=1}^n d\lambda_i \Delta(\lambda_k) f(\lambda) p(\lambda) \right) \end{aligned}$$

Таким образом, мы выяснили, что якобиан J в точности равен $\Delta(H)$. В нашей «умной единице» мы выделили явно меру $\int [DU]$, чтобы потом, сделав замену переменной и убрав зависимость от неё в подинтегральном выражении, вынести её за скобку и сократить. Такой трюк выделения в виде отдельного фактора интеграла по лишним переменным, от которых не зависят усредняемые выражения, был изобретен в калибровочных теориях поля.



Упражнение. В рассказе выше был небольшой обман. Подумайте, в чём он заключался и почему он не влияет существенно на итог выкладок.

Ещё раз повторим результат: среднее значение от величины, зависящей только от собственных значений H , можно вычислять с помощью N -мерного интеграла

$$\prod_{i=1}^N \int d\lambda_i \times \prod_{i < j} (\lambda_i - \lambda_j)^2 \exp \left[-N \sum_{i=1}^N V(\lambda_i) \right] \times f(\lambda)$$

В заключение рассмотрим, как применить это для вычисления лидирующего ответа для плотности $\rho(E)$. В пределе больших N этот многомерный интеграл содержит большой параметр в экспоненте и его можно взять по *методу перевала*; напомним, что в одномерном случае он говорит, что для большого параметра λ , функции f с минимумом в точке x_0 и медленно меняющейся в окрестности x_0 функции g

$$\int dx \exp(-\lambda f(x)) g(x) \approx \frac{1}{\sqrt{\lambda |f''(x_0)|}} g(x_0) \exp(-\lambda f(x_0))$$

(для нашего случая префактор не будет важен). Перевальная точка имеет смысл «наиболее вероятной конфигурации» в пределе $\lambda \rightarrow \infty$: в нашем случае распределение λ_i , дающих минимум, логично считать приближением нулевого порядка по $1/N$ для средней плотности собственных значений.

Перевальная точка даётся минимизацией по всем λ_j выражения вида

$$\sum_{i \neq j} \log |\lambda_i - \lambda_j| - N \sum V(\lambda_i)$$

что приводит к N уравнениям

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

Предполагая, что λ в пределе распределены квазинепрерывно, сумму по j мы можем заменить на интеграл с искомой плотностью ρ , а уравнение распространить на любую точку x , которая лежит в области, где $\rho \neq 0$. Тогда оно примет вид

$$V'(x) = 2 \int d\lambda \left(\mathcal{P} \frac{1}{x - \lambda} \right) \rho(\lambda)$$



Упражнение. Вычислите интеграл, подставив плотность в виде «вигнеровского полукруга» и убедитесь, что он даёт квадратичный потенциал $V(x) = \frac{x^2}{2}$.

Используя некоторые соображения из комплексного анализа, это уравнение на самом деле можно решить точно для любого полиномиального потенциала. Это, однако, выходит за рамки нашего рассказа. Спасибо всем, кто дочитал :)

5. Литература

- 1) Mehta, «Random matrices»: большая подробная книжка с кучей методов и точных результатов для различных (преимущественно гауссовых) ансамблей
- 2) Makeenko, «Methods of contemporary gauge theory»: сжато изложены основы матричных моделей, large-N методы и связь с теорией поля
- 3) Eynard, Kimura, Ribault «Random matrices»: обзор для математически ориентированных читателей, включает много продвинутых тем типа топологической рекурсии и петлевых уравнений