World Scientific
www.worldscientific.com

COLLECTIVE POTENTIAL FOR LARGE-N HAMILTONIAN MATRIX MODELS AND FREE FISHER INFORMATION

A. AGARWAL,[*] L. AKANT,[†] G. S. KRISHNASWAMI[‡] and S. G. RAJEEV[§]

*University of Rochester, Department of Physics and Astronomy,
Rochester, New York, 14627, USA*
[*] *abhishek@pas.rochester.edu*
[†] *akant@pas.rochester.edu*
[‡] *govind@pas.rochester.edu*
[§] *rajeev@pas.rochester.edu*

We formulate the planar "large N limit" of matrix models with a continuously infinite number of matrices directly in terms of $U(N)$ invariant variables. Noncommutative probability theory is found to be a good language to describe this formulation. The change of variables from matrix elements to invariants induces an extra term in the Hamiltonian, which is crucial in determining the ground state. We find that this collective potential has a natural meaning in terms of noncommutative probability theory: it is the "free Fisher information" discovered by Voiculescu. This formulation allows us to find a variational principle for the classical theory described by such large N limits. We then use the variational principle to study models more complex than the one describing the quantum mechanics of a single Hermitian matrix (i.e. go beyond the so-called $D = 1$ barrier). We carry out approximate variational calculations for a few models and find excellent agreement with known results where such comparisons are possible. We also discover a lower bound for the ground state by using the noncommutative analog of the Cramer–Rao inequality.

Keywords: Matrix model; Yang–Mills theory; large N limit; collective potential; variational principle; noncommutative probability theory; free Fisher information; free entropy.

1. Introduction

The purpose of the present paper is to find the natural mathematical language for describing the "large N limit" of matrix models involving a continuously infinite number of matrices (in the Hamiltonian picture) e.g. matrix quantum mechanics and matrix field theories. We find that noncommutative probability theory offers such a language, which after suitable generalizations can be adopted for the purposes of formulating the large N limit. Recasting continuum matrix models in this language allows us to find a variational principle for calculating the gauge invariant observables of such models. The advantage of the variational method being that

we can use it to make well-controlled approximations when exact solutions are no longer possible.

No apologies are really necessary for the study of continuum matrix models as they are ubiquitous in the world of both gauge theories and strings. QCD, where the gluonic field (A_μ) belongs to the adjoint representation of the structure group, is perhaps the most natural example of a field theory where the dynamical degrees of freedom are matrix valued. It is believed that the planar large N limit is the correct simplification to consider while trying to understand nonperturbative phenomena such as confinement, understanding which in the physically interesting case of four space–time dimensions remains one of the major unsolved problems of theoretical physics. QCD when dimensionally reduced to one dimension is known to provide a regularization of the theory of the classical membrane, and the generalization of this method to the case of the supermembrane provides one with the Hamiltonian which is conjectured to be a description of M-theory in the infinite momentum frame.[1] Other examples of one-dimensional matrix models include those with a single matrix (with polynomial self interactions) depending continuously on a parameter (which is usually taken to be physical or stochastic time). Such models are known to provide explicit realizations of string field theories in one-dimensional target spaces.[2,3]

In a previous paper[8] we studied "static" matrix models, involving the integral over a finite number of Hermitian matrices:

$$\int e^{N \operatorname{tr} S(A)} \frac{1}{N} \operatorname{tr} A_{i_1} \cdots A_{i_n} dA \,. \tag{1}$$

It is known that the theory of "free random variables"[4] has a lot to do with the large N limit of these models.[5−8] The mathematician's notion of "freeness"[9−12] and the physicist's picture of the planar large N limit are intimately related. The large N limit of matrix models was formulated in terms of a variational principle. There is an anomalous term in this variational principle, which we showed, has a simple meaning in terms of noncommutative probability theory: it is the free entropy of Voiculescu.

In the present paper we will extend our study to "dynamical" matrix models, determined by Hamiltonians such as

$$-K_{ij} \frac{\partial^2}{\partial A_{ia}^b \partial A_{jb}^a} + \operatorname{tr} \tilde{V}(A) \,. \tag{2}$$

In the path integral formulation of quantum mechanics this leads to integrals such as

$$\int e^{- \int [K_{ij} \operatorname{tr} \dot{A}_i \dot{A}_j + \tilde{V}(A)] dt} \operatorname{tr} A_{i_1}(t_1) \cdots A_{i_n}(t_n) \mathcal{D} A \,. \tag{3}$$

Starting from the quantum theory at finite "N," there has been considerable progress in extracting the large N limit, which is described by a classical Hamiltonian. The "Collective Field" formalism of Jevicki and Sakita[13−17] and the "coherent state" picture of Yaffe[18] stand out in this area. The idea is to change variables from

the matrix elements to the invariants of the matrices, because quantum fluctuations in these invariants become small in the large N limit. Thus the theory when expressed in terms of invariants should be a classical theory. One of the key lessons of their work is that this large N classical limit differs in many deep ways from the more familiar classical limit as $\hbar \to 0$. Most striking is the appearance of certain universal terms in the gauge invariant formulation of the large N theory (e.g. The ϕ^3 term in the collective potential of Jevicki and Sakita).

To understand the origin of these terms, let us consider a toy model from elementary quantum mechanics: a rotation invariant system whose configuration space is R^N and has Hamiltonian operator

$$\tilde{H} = -\hbar^2 \frac{\partial^2}{\partial x^i \partial x^i} + \tilde{V}(r), \quad r = \sqrt{[x^i x^i]}. \tag{4}$$

If we consider only wave functions that are rotation invariant (zero angular momentum), the angular variables can be eliminated and the Hamiltonian can be expressed as a differential operator in the invariant variable r. The simplest way to perform this change of variables is to note that

$$\int [\hbar^2 \partial_i \psi^* \partial_i \psi + \tilde{V}(r)|\psi|^2] d^N x = C_N \int [\hbar^2 |\psi'|^2 + \tilde{V}(r)|\psi|^2] r^{N-1} dr, \tag{5}$$

where $C_N r^{N-1}$ is the area of the sphere of radius r. This factor is the Jacobian of the nonlinear change of variables $x^i \to r$. Even if we make the change of variables to the "radial wave function" $\Psi(r) = r^{\frac{N-1}{2}} \psi(r)$ to absorb this Jacobian factor, there will be some derivatives of the Jacobian that survive. After an integration by parts, we get

$$C_N \int \left[\hbar^2 |\Psi'|^2 + \left\{ \hbar^2 \frac{(N-1)(N-3)}{4r^2} + \tilde{V}(r) \right\} |\Psi|^2 \right] dr. \tag{6}$$

Thus, the radial Hamiltonian is not just $-\hbar^2 \frac{\partial^2}{\partial r^2} + V(r)$, but is instead,

$$\tilde{H} = -\hbar^2 \frac{\partial^2}{\partial r^2} + \hbar^2 \frac{(N-1)(N-3)}{4r^2} + \tilde{V}(r). \tag{7}$$

The extra term is a sort of "centrifugal potential" arising from differentiation of the volume of the sphere of radius r. It is at first surprising that such a centrifugal potential is present even when the angular momentum is zero: a phenomenon with no analog in the classical theory. Indeed, this term vanishes as $\hbar \to 0$. This term is universal in the sense that it is independent of the choice of potential: it arises from rewriting the kinetic energy in radial variables.

To complete the story of this toy model, we see how to take the large N limit. In this limit we should expect that

$$\rho^2 = \frac{1}{N} x_i x_i \tag{8}$$

should have small fluctuations. (This is a kind of central limit theorem.) Thus expressed in these variables, the theory should approach a classical limit. This

classical limit will be very different from the more obvious limit $\hbar \to 0$. In particular, the ground state can be very different. The Hamiltonian is

$$\tilde{H} = -\frac{\hbar^2}{N}\frac{\partial^2}{\partial\rho^2} + \hbar^2\frac{(N-1)(N-3)}{4N\rho^2} + \tilde{V}(N^{\frac{1}{2}}\rho)\,. \tag{9}$$

If we define $H = \frac{\tilde{H}}{N}$, $V(r) = \frac{1}{N}\tilde{V}(N^{\frac{1}{2}}\rho)$ we get

$$H = -\frac{\hbar^2}{N^2}\frac{\partial^2}{\partial\rho^2} + \hbar^2\frac{(N-1)(N-3)}{4N^2\rho^2} + V(\rho)\,. \tag{10}$$

If we define

$$\pi = -\frac{i}{N}\frac{\partial}{\partial\rho} \tag{11}$$

we get the commutation relations

$$[\pi,\rho] = -\frac{i}{N} \tag{12}$$

and the Hamiltonian

$$H = -\hbar^2\pi^2 + V_{\text{coll}}(\rho) + V(\rho)\,, \tag{13}$$

where $V_{\text{coll}} = \frac{\hbar^2(N-1)(N-3)}{N^2}\frac{1}{4\rho^2}$. Now we can take the limit as $N \to \infty$, keeping $V(\rho)$, $\hbar$ fixed. The collective potential tends to

$$V_{\text{coll}}(\rho) \to \frac{\hbar^2}{4}\Phi(\rho)\,, \quad \Phi(\rho) = \frac{1}{\rho^2}\,. \tag{14}$$

The quantum fluctuations in ρ are indeed small because the commutator $[\pi,\rho] \sim \frac{1}{N}$. In the limit we get a classical Hamiltonian

$$H = \hbar^2p^2 + \frac{\hbar^2}{4}\Phi(\rho) + V(\rho) \tag{15}$$

with the canonical Poisson bracket relations

$$\{\rho,p\} = 1\,. \tag{16}$$

The ground state of this system is given by $p = 0$, $\rho = \rho_0$ where ρ_0 is the minimum of

$$\frac{\hbar^2}{4}\Phi(\rho) + V(\rho)\,. \tag{17}$$

In the conventional classical limit $\hbar \to 0$ keeping N fixed, the ground state would have been at the minimum of $V(\rho)$. For example, if $V(\rho) = \frac{1}{2}\omega^2\rho^2$, it has the minimum at $\rho = 0$ with vanishing ground state energy. On the other hand, the ground state of the large N limit is concentrated at

$$\rho_0 = \left[\frac{\hbar}{\sqrt{2}\omega}\right]^{\frac{1}{2}} \tag{18}$$

with $E_{\text{gs}} = \frac{\hbar\omega}{\sqrt{2}}$. Thus the large N limit better captures the zero point fluctuations of this system.

The passage from matrix elements to invariants of a collection of matrices is analogous to the above change of variables from Cartesian to radial coordinates. There is a Jacobian associated with this nonlinear change. Transforming to "radial wave functions" to absorb this will again induce an extra term in the potential, involving the derivative of the Jacobian. These universal terms enable the large N limit to capture features of the theory that the usual perturbation expansion misses. They are crucial in determining the ground state of the matrix model. Is it possible to understand them directly in the large N limit?

In this paper we show that this universal term has a natural meaning in the world of noncommutative probability theory: it is the free analog of the Fisher information obtained by Voiculescu.[9−12] This continues our earlier work[8] where it is shown that the Jacobian of the change of variables to invariants is just the exponential of free entropy. Information is the square of the gradient of entropy, which gives a rough explanation of our present result.

We first motivate the connection between free probability theory and large N matrix models in the context of static models, explaining some concepts from noncommutative probability theory that appear not to be widely known in the physics literature. Then we carry out this reformulation for the problem of matrix quantum mechanics (or field theory at D (dimension of space–time) $= 1$). This allows us to recover the results of Jevicki–Sakita.[13−17] Certain points of view on noncommutative probability theory generalize beyond $D = 1$, and we use them to formulate a variational principle for general matrix models. In the last part of the paper, we apply this principle to some specific examples to get approximate answers for the ground state of matrix field theories with $D > 1$, where explicit solutions are not possible.

2. Noncommutative Probability and Static Matrix Models

Recall[4,8] that in noncommutative probability theory, the analog of a probability distribution is a rule that gives an expectation value of a polynomial in the basic random variables. More precisely, consider the free associative algebra $\mathcal{T}_M = C\langle X_1, \ldots, X_M \rangle$ generated by the Hermitian operator valued random variables X_i over the complex numbers. We think of the elements of $\mathcal{T}_M$ as formal power series[a] $f(X) = f^{i_1 \cdots i_n} X_{i_1 \cdots i_n} = f^I X_I$. The joint probability distribution of the X_i in a tracial state is specified by the expectation values $G_I = \langle X_I \rangle$, which are cyclically symmetric tensors ("moments") satisfying

$$G_\phi = 1, \quad G_I^* = G_{\bar{I}}, \quad G_{IJ} f^{\bar{I}^*} f^J \geq 0 \quad \text{for any polynomial } f(X), \qquad (19)$$

where $\bar{I}$ is the string with the opposite order. The last condition is the positivity of the "Hankel matrix" $G_{I;J} = G_{IJ}$. It is used to construct a metric on the space of probability distributions. $G^{I;J} G_{JK} = \delta_K^I$ denotes the inverse.

---

[a]We continue the notation in Ref. 8. An upper case letter such as I denotes a string of indices $i_1 \cdots i_n$. Repeated multi-indices are summed; ϕ denotes the empty string.

If there were only one such random matrix X (i.e. $M = 1$), it would be possible to introduce a probability density function on the real line $\rho_G(x)$ such that

$$G_n = \int \rho_G(x) x^n dx \, . \tag{20}$$

Many interesting quantities (such as entropy or information) have simple formulas in terms of $\rho(x)$.

The main complication in passing to multimatrix models with $M > 1$ is that there is no longer any analog of $\rho(x)$. We have to think of noncommutative probability distributions indirectly in terms of the moments G_I or certain other cyclically symmetric tensors S^I which contain equivalent data.

The two sets of tensors S^I and G_I are related by the "factorized Schwinger–Dyson equations:"

$$S^{J_1 i J_2} G_{J_1 I J_2} + \eta_I^i = 0 \, , \tag{21}$$

where $\eta_I^i = \delta_I^{I_1 i I_2} G_{I_1} G_{I_2}$.

The motivation behind this definition is that the tensors S^I can be thought of as providing a "matrix model" for the noncommutative probability distribution:

$$G_I = \lim_{N \to \infty} \frac{\int dA e^{N \operatorname{tr} S(A)} \frac{\operatorname{tr}}{N} A_I}{\int dA e^{N \operatorname{tr} S(A)}} \, . \tag{22}$$

Thus, we can think of $\frac{e^{N \operatorname{tr} S(A)}}{\int dA e^{N \operatorname{tr} S(A)}}$ as a (commutative) probability density function for the matrix elements of a set of $N \times N$ matrices A_i. The quantity $\operatorname{tr} S(A)$ is called the "action" of the matrix model in the physics literature. This extends to the realm of random variables the usual idea that noncommutative variables can be realized as matrices with complex numbers as entries. In a typical problem arising from quantum physics, the action $S(A)$ is some simple polynomial and we are required to calculate the moments arising from it: i.e. solve the Schwinger–Dyson equations.

Infinitesimal variations of a noncommutative probability distribution are vector fields L_I^i on the space of probability distributions, with a natural action on the moments given by,

$$L_I^i G_J = \delta_J^{J_1 i J_2} G_{J_1 I J_2} \, . \tag{23}$$

Algebraically, L_I^i are the derivations of the tensor algebra $\mathcal{T}_M$ corresponding to the infinitesimal changes of variables $[\delta A]_j = \delta_j^i A_I$. These vector fields form a Lie algebra, with the following Lie bracket:

$$[L_I^i, L_J^j] = \delta_J^{J_1 i J_2} L_{J_1 I J_2}^j - \delta_I^{I_1 j I_2} L_{I_1 J I_2}^i \, . \tag{24}$$

For the case of the one matrix model, when the strings are labeled by integers, this algebra reduces to the Virasoro algebra (without the central term).

The quantity η_I^i appearing in the Schwinger–Dyson equation generates the first cohomology of the Lie algebra:[8,20] it can be viewed as the variation $\eta_I^i(G) = L_I^i \chi(G)$ of some function $\chi(G)$ which, however cannot be expressed as a formal power series

in the moments. Thus it is possible to regard the Schwinger–Dyson equation as the condition for a function

$$\Omega(G) = S^I G_I + \chi(G) \tag{25}$$

to be stationary: $L_I^i \Omega = 0$. This variational principle is useful because it allows for approximate solution of the SD equations: it is rare that they can be solved exactly.

This quantity $\chi(G)$ has a natural meaning in noncommutative probability theory: up to an additive constant, $\chi(G)$ is the free entropy in the sense of Voiculescu.[9] It is the logarithm of volume of the set of $N \times N$ Hermitian matrices whose moments are G_I in the $N \to \infty$ limit.[21]

To obtain a more explicit formula for χ, we have to introduce a third way of describing a noncommutative probability distribution: a transformation $\phi : A_i \mapsto \phi(A)_i = \phi_i^I A_I$ that maps given distribution to a reference distribution, i.e. $\phi : G \to \Gamma$. Under such a change of variable, χ transforms as a one-cocycle, it is incremented by the expectation value of $\frac{1}{N^2} \log \det$ of the Jacobian of ϕ. We do not follow this direction here, as it is explained in an earlier paper.[8]

For a single Hermitian random variable, there are simple explicit formulas for entropy

$$\chi(G) = \mathcal{P} \int dx\, dy\, \rho_G(x) \rho_G(y) \log |x - y|$$

$$= \int dx\, dy\, \rho_\Gamma(x) \rho_\Gamma(y) \log \left(\frac{\phi(x) - \phi(y)}{x - y} \right) + \chi(\Gamma), \tag{26}$$

where $\rho_\Gamma(x) = \rho_G(\phi(x)) \phi'(x)$.

It is also worth recalling, that for the case of a single random matrix, the Schwinger–Dyson equation can be written in a simpler form using ρ:

$$(H\rho)(x) = \frac{1}{2\pi} S'(x). \tag{27}$$

(This is called the Mehta–Dyson integral equation.) Here $(H\rho)(x) = \frac{1}{\pi} \mathcal{P} \int dy \frac{\rho(y)}{x - y}$ is the Hilbert transform.

The expression $\eta_I^i = \delta_I^{I_1 i I_2} G_{I_1} G_{I_2}$ motivates the introduction of the operator ∂,[10] the difference quotient gradient, whose components are

$$\partial^j : \mathcal{T}_M \to \mathcal{T}_M \otimes \mathcal{T}_M; \quad \partial^j X_I = \delta_I^{I_1 j I_2} X_{I_1} \otimes X_{I_2}. \tag{28}$$

The partial difference quotient can be thought of as acting on a product of matrices. It breaks the matrix product at each occurrence of X_j, exposing two more matrix indices, thus producing an element in the tensor product. For $M = 1$, $\partial \phi(x) = \frac{\phi(x) - \phi(y)}{x - y}$. ∂ relates the noncommutative Hilbert transform to the variation of entropy.

The analog of the difference quotient, when acting on a trace of a product of matrices rather than on a product of matrices, leads to the cyclic gradient δ:[11]

$$\delta^i : \mathcal{T}_M \to \mathcal{T}_M; \quad \delta^i X_I = \delta_I^{I_1 i I_2} X_{I_2 I_1}. \tag{29}$$

The ordering of multi-indices is explained by the matrix analogy. For a single generator, the cyclic gradient reduces to the partial derivative. The cyclic gradient of a polynomial is unchanged by cyclic permutations of any of its coefficients: Let $S(X) = S^I X_I$. Then

$$\delta^k S(X) = S^{I_1 k I_2} X_{I_2 I_1} = S^{(kK)} X_K \,, \tag{30}$$

where $S^{(K)} = S^{k_1 \cdots k_m} + S^{k_m k_1 \cdots k_{m-1}} + \cdots + S^{k_2 \cdots k_m k_1}$.

The formalism developed above allows us to get a formula for the Hilbert transform of a collection of noncommutative random variables. We first note that for a single Hermitian variable, the Mehta–Dyson equation $H\rho(x) = \frac{1}{2\pi} S'(x)$ tells us that the Hilbert transform is the derivative of the one-matrix model action. For several Hermitian random variables, the Hilbert transform is a vector $H^i(X)$. The appropriate generalization of the "derivative" of the action $S^I A_I$, must depend only on the cyclic part of the tensors S^I. The natural candidate is the cyclic gradient

$$H^i(X) = \frac{1}{2\pi} \delta^i S(X) \,. \tag{31}$$

Using the definition of δ^i,

$$H^i(X) = \frac{1}{2\pi} S^{K_1 i K_2} X_{K_2 K_1} = \frac{1}{2\pi} S^{Ii} X_I \tag{32}$$

the formula for the Hilbert transform given above involves the action $S(X)$. One can look for a formula for $H^i(X)$ in terms of the moments of the matrix model. Using the factorized SD equations in the form $S^{Ii} = \eta^i_J G^{I;J}$ we get

$$H^i(X) = \frac{1}{2\pi} \eta^i_I G^{I;J} X_J = \frac{1}{2\pi} G_{I_1} G_{I_2} G^{I_1 i I_2; J} X_J \,. \tag{33}$$

Obtaining the inverse Hankel matrix from the moments is about as hard as getting the moments of a static matrix model given its action. This formula may be written

$$L^i_I \chi = 2\pi \langle H^i(X) X_I \rangle \tag{34}$$

thereby relating the Hilbert transform to the variation of entropy. Voiculescu used this as the defining property of the Hilbert transform.[10]

Another quantity from the world of noncommutative probability theory that is going to be important for our understanding of matrix models is the notion of free Fisher information. Following Voiculescu, let us first consider Fisher information for a single real random variable f:

$$\Phi(f) = \lim_{t \to 0^+} \frac{\mathcal{S}(f + \sqrt{t}g) - \mathcal{S}(f)}{t} \,. \tag{35}$$

Here g is Gaussian with unit covariance, f and g are statistically independent and $\mathcal{S}(f)$ is Boltzmann's entropy of f. If f has pdf $\rho(x)$, then

$$\mathcal{S}(f) = - \int \rho(x) \log \rho(x) dx \,. \tag{36}$$

Regarding Φ as the zero time derivative of a Gaussian process starting at f, the heat equation gives

$$\Phi(f) = \int dx\, \rho(x)(\partial_x \log \rho(x))^2 \,. \tag{37}$$

For several real random variables x_i with joint distribution $\rho(\mathbf{x})$, the Fisher information matrix is defined as

$$\Phi^{ij} = \int d\mathbf{x}\, \rho(\mathbf{x})\partial^i \log \rho(\mathbf{x})\partial^j \log \rho(\mathbf{x}) \,. \tag{38}$$

By analogy, the free Fisher information of a single Hermitian random variable may be defined as:[10]

$$\Phi(X) = \lim_{t \to 0^+} \frac{\chi(X + \sqrt{t}S) - \chi(X)}{t} \,, \tag{39}$$

where S is the Wigner semicircular distribution. If $X + \sqrt{t}S$ has pdf $\rho(x,t)$, then $\chi(X + \sqrt{t}S) = \mathcal{P} \int \rho(x,t)\rho(y,t) \log|x - y| dx\, dy$ is the free entropy. The complex Burger equation

$$\frac{\partial G}{\partial t} + G\frac{\partial G}{\partial z} = 0 \tag{40}$$

for the Cauchy transform $G(z,t) = \int \frac{\rho(x,t)dx}{x-z}$, $z \notin \mathrm{supp}(\rho)$ plays the role of the heat equation for a semicircular process and leads to

$$\Phi(X) = 2\pi^2 \int ((H\rho)(x))^2 \rho(x) dx = \frac{2\pi^2}{3} \int dx(\rho(x))^3 \,. \tag{41}$$

By analogy with the case of several real random variables, we define the free Fisher information matrix for several noncommuting random variables as

$$\Phi^{ij}(G) = 2\pi^2 \langle H^i(X)H^j(X) \rangle \,, \tag{42}$$

$\Phi^{ij}(G)$ cannot be expressed solely in terms of the moments. As with the Hilbert transform, we can express it in terms of the G_K and either the inverse Hankel matrix $G^{I;J}$ or the action $S^I X_I$ whose large N moments are G_K:

$$\Phi^{ij}(G) = \frac{1}{2}G^{I;J}\eta_I^i \eta_J^j$$

$$= \frac{1}{2}G^{I_1 i I_2; J_1 j J_2} G_{I_1} G_{I_2} G_{J_1} G_{J_2} \tag{43}$$

$$\Phi^{ij}(G) = \frac{1}{2}S^{iI}S^{jJ}G_{IJ} \,.$$

The equivalence of these two follows from the factorized SD equations relating S^I to G_K:

$$S^{Ii} = \eta_J^i G^{I;J} \,. \tag{44}$$

For a single generator this formula reduces to the familiar one when we use the Mehta–Dyson equation.

$$\Phi(G) = \frac{1}{2} \sum_{m,n} (m+1)(n+1) S_{m+1} S_{n+1} G_{m+n}$$

$$= \frac{1}{2} \langle (S'(x))^2 \rangle = 2\pi^2 \int ((H\rho)(x))^2 \rho(x) dx \tag{45}$$

3. Dynamical Matrix Models and Collective Field Theory

Now let us turn to the main object of our interest, i.e. matrix models in the Hamiltonian approach. To have a definite example in mind we consider in some detail, the quantum mechanics of a discrete number of Hermitian matrices. In other words, we consider dynamical multimatrix models whose Hamiltonians are of the form

$$\tilde{H} = -K_{ij} \frac{1}{2} \frac{\partial^2}{\partial(A_i)_b^a \partial(A_j)_a^b} + \mathrm{tr}\left[\tilde{V}^I A_I\right]. \tag{46}$$

Here $A_i(t)$ are $N \times N$ Hermitian matrices, whose time evolution is described by the Hamiltonian above. The positive tensor K_{ij} determining the kinetic energy is usually just $\hbar^2 \delta_{ij}$. The ground state energy is

$$E_{\mathrm{gs}} = \lim_{N \to \infty} \frac{1}{N^2} \min_{\|\psi(A)\|=1} \langle \psi | \tilde{H} | \psi \rangle, \tag{47}$$

where the wave function $\psi(A)$ is invariant under the adjoint action of the unitary group; $A_i \mapsto U A_i U^\dagger \forall i$ and the inner product is with respect to the Lebesgue measure on matrix elements.

The planar large N limit of the above theory is the limit $N \to \infty$ holding $V^I = N^{|I|-1} \tilde{V}^I$ and K_{ij} fixed. Moreover, only wave functions that are invariant under $U(N)$ are allowed. It will be convenient to make the rescaling $A_i \to N A_i$ so that

$$\frac{\tilde{H}}{N^2} = H = -\frac{1}{N^4} K_{ij} \frac{1}{2} \frac{\partial^2}{\partial(A_i)_b^a \partial(A_j)_a^b} + \frac{1}{N} \mathrm{tr}\left[V^I A_I\right]. \tag{48}$$

It is known that in this limit, quantum fluctuations in the moments $G_I = \frac{1}{N} \mathrm{tr}\, A_I$ become small. Hence there ought to be some classical mechanical system that is equivalent to this limiting case of the matrix model. The configuration space of this classical mechanical system would be the space of probability distributions. It is our task now to determine the Hamiltonian and Poisson brackets of this theory and use them to gain insight into the properties (such as the ground state) of the matrix model.

3.1. *The case of one degree of freedom*

Such a reformulation is most transparent in the case where the time evolution of a single Hermitian matrix is concerned,

$$\tilde{H} = -\frac{1}{2}\hbar^2 \frac{\partial^2}{\partial A_b^a \partial A_a^b} + \sum_n \operatorname{tr} \tilde{V}_n A^n \,. \tag{49}$$

If we rescale $A \to NA$ this becomes

$$\frac{\tilde{H}}{N^2} = -\frac{1}{2}\frac{\hbar^2}{N^4}\frac{\partial^2}{\partial A_b^a \partial A_a^b} + \sum_n V_n \frac{1}{N}\operatorname{tr} A^n \,, \tag{50}$$

where $V_n = N^{n-1}\tilde{V}_n$. We are interested in determining the properties (such as the ground state) of the system in the limit as $N \to \infty$ keeping V_n and $\hbar$ fixed.

We quote below the result obtained by Jevicki and Sakita,[14,16,17] who changed variables to the eigenvalue density $\rho(x)$ of A and showed that in the large N limit, the Hamiltonian may be written as

$$\frac{\tilde{H}}{N^2} = H = K_r + V_{\text{coll}} + \int dx\, V(x)\rho(x) \,,$$

where

$$V(x) = \sum_n V_n x^n \,. \tag{51}$$

The collective potential and "radial" kinetic energy[b] are

$$V_{\text{coll}} = \hbar^2 \frac{\pi^2}{6} \int dx\, \rho(x)^3 = \hbar^2 \frac{\pi^2}{2} \int ((H\rho)(x))^2 \rho(x)dx \,,$$

$$K_r = \frac{1}{2}\hbar^2 \int dx\, \pi'(x)\rho(x)\pi'(x) \,,$$

where

$$\pi(x) = -\frac{i}{N^2}\frac{\delta}{\delta\rho(x)} \,; \quad [\pi(x), \rho(x')] = -\frac{i}{N^2}\delta(x - x') \,. \tag{52}$$

We notice that their collective potential V_{coll} (the term that really makes the Hamiltonian describing the large N limit different from the finite N case) is just one fourth the free Fisher information of one Hermitian random variable.

This reformulation of the matrix model allows us to take the limit $N \to \infty$ easily. Note that N only appears in the commutation relations between ρ and π. Indeed, as $N \to \infty$ this commutator becomes small: $\frac{1}{N^2}$ appears where $\hbar$ appears in the Heisenberg relations. Thus $N \to \infty$ is a kind of classical limit. The limiting theory has as configuration space the space of probability distributions $\rho(x)$. The canonical conjugate of $\rho(x)$ is $p(x)$:

$$\{\rho(x), p(x')\} = \delta(x - x') \,. \tag{53}$$

---

[b]We use the term radial by analogy with atomic physics.

The Hamiltonian

$$H = \frac{1}{2} \int dx\, \hbar^2 p'^2(x)\rho(x) + V_{\text{coll}}(\rho) + \int V(x)\rho(x)dx \tag{54}$$

leads to the equations of motion

$$\frac{\partial p(x,t)}{\partial t} = E - \frac{1}{2}\hbar^2 p'^2(x) - \frac{\pi^2 \hbar^2}{2}\rho(x)^2 - V(x)\,,$$
$$\frac{\partial \rho(x,t)}{\partial t} = -\hbar^2 \frac{\partial[\rho(x)p'(x)]}{\partial x}\,, \tag{55}$$

where E is the Lagrange multiplier imposing the constraint $\int \rho(x)dx = 1$. The spectrum (density of eigenvalues) of the matrix A in ground state of the matrix quantum mechanics is then determined in this large N approximation by the static solution

$$\rho(x) = \frac{1}{\pi \hbar}\sqrt{[2(E - V(x))]}\theta(E - V(x))\,. \tag{56}$$

As noted above, the constant E is determined by the condition $\int \rho(x)dx = 1$.

Notice that the collective potential V_{coll} plays a crucial role in determining this answer; without that term there would have been no such static (ground state) solution. The elimination of the angular degrees of freedom induces a centrifugal potential (collective potential V_{coll}) leading to a repulsion of eigenvalues of A. The volume of the set of matrices of a given spectrum vanishes when any pair of eigenvalues coincide. The collective potential is the square of the derivative of the logarithm of this volume, so is singular when a pair of eigenvalues coincide. This repulsion of eigenvalues counterbalances the tendency of the eigenvalues to accumulate at the minimum of $V(x)$. It is this balance that leads to a continuous density of eigenvalues.

In the more conventional limit $\hbar \to 0$ (keeping N fixed) we would get a completely different answer: the ground state is concentrated at the minimum of $V(x)$.

3.2. *Several degrees of freedom*

Once we see the connection of the collective potential to the free Fisher information, noncommutative probability theory immediately suggests the generalization to multimatrix models. (We will describe the answer first, postponing its derivation to the next subsection.) The probability density $\rho(x)$ no longer makes sense with several random variables; instead we must think of the moment tensors G_I as characterizing the noncommutative probability distribution. This way of thinking leads to

$$H = \frac{1}{2}K_{ij}\pi^{iI}G_{IJ}\pi^{jJ} + \frac{1}{4}K_{ij}\Phi^{ij}(G) + V^I G_I\,,$$
$$\pi^I = -\frac{i}{N^2}\frac{\partial}{\partial G_I}\,; \qquad [\pi^I, G_J] = -\frac{i}{N^2}\delta^{(J)}_{(I)}\,, \tag{57}$$

where $\delta_{(I)}^{(J)} = 1$ if there is a cyclic permutation of I that equals J and zero otherwise. We will now show by an explicit change of variables that this is indeed the correct Hamiltonian of the large N limit. It is convenient to introduce the effective potential

$$V_{\text{eff}} = \frac{1}{4} K_{ij} \Phi^{ij} + V^I G_I. \tag{58}$$

Thus, in the planar limit $N \to \infty$ keeping V^I, K_{ij} fixed we do get a classical mechanical system whose configuration space is the space of noncommutative probability distributions with Poisson brackets

$$\{G_I, \mathbf{p}^J\} = \delta_{(I)}^{(J)} \tag{59}$$

and the Hamiltonian

$$H = \frac{1}{2} K_{ij} p^{iI} G_{IJ} p^{jJ} + \frac{1}{4} K_{ij} \Phi^{ij}(G) + V^I G_I \tag{60}$$

determine the equations of motion

$$\frac{dG_K}{dt} = \{G_K, H\} = \{G_K, K_r\}$$

$$= \frac{1}{2} K_{ij} \delta_{(K)}^{(J_1 j J_2)} \left[\pi^{I_1 i I_2} G_{I_2 I_1 J_2 J_1} + G_{I_2 I_1 J_2 J_1} \pi^{I_1 i I_2} \right], \tag{61}$$

$$\frac{d\pi^K}{dt} = \{\pi^K, H\} = \frac{1}{2} K_{ij} \delta_{(I_2 I_1 J_2 J_1)}^{(K)} \pi^{I_1 i I_2} \pi^{J_1 j J_2} + \frac{\partial V_{\text{eff}}(G)}{\partial G_K}. $$

An interesting class of static solutions is determined by

$$\pi^K = 0 \ \forall \, |K| \geq 1 \,,$$
$$\frac{\partial V_{\text{eff}}(G)}{\partial G_K} = 0 \ \forall \, |K| \geq 1 \,. \tag{62}$$

Thus, to find the ground state we must minimize V_{eff}. Note that this variational principle involves no approximation aside from the large N limit. The equations for extremizing V_{eff} are, $Lk_K \left(\frac{1}{4} K_{ij} \Phi^{ij} + V^I G_I \right) = 0$. The chief obstacle in solving these equations explicitly is the presence of the Fisher information term, which does not have any simple expression in terms of the moments. Of the two forms of the information matrix in Eq. (43), we choose to use the one involving the auxiliary variables S^I's in carrying out the minimization of the effective potential. Let us note here, that these variables $S^I(G)$ are to be thought of as functions of the moments themselves, i.e. they are the solution to the Schwinger–Dyson equations of some (static) matrix model, whose large N moments coincide with those of the dynamical matrix model at a particular value of time. In terms of the auxiliary variables $S^I(G)$ the equations to be solved are

$$L_K^k V_{\text{eff}} = \frac{1}{8} K_{ij} S^{jJ} \left[2 L_K^k \eta_J^i - S^{iI} L_K^k G_{IJ} \right] + V^{I_1 k I_2} G_{I_1 K I_2} = 0 \,. \tag{63}$$

3.3. *A geometric derivation of the collective Hamiltonian*

We will study the quantum mechanical system whose degrees of freedom are described by a set of Hermitian matrices $(A_i)^a_b$ and has Hamiltonian

$$H = -\frac{1}{N^4} K_{ij} \frac{1}{2} \frac{\partial^2}{\partial (A_i)^a_b \partial (A_j)^b_a} + \frac{1}{N} \operatorname{tr}[V^I A_I]. \tag{64}$$

The eigenvalue problem of H follows from the condition that the quadratic form

$$\langle \psi | H | \psi \rangle = \int \left[\frac{1}{2N^4} K_{ij} \frac{\partial \psi^*}{\partial A^a_{ib}} \frac{\partial \psi}{\partial A^b_{ja}} + \frac{1}{N} \operatorname{tr} V^I A_I |\psi(A)|^2 \right] \prod_{i=1}^{M} d^{N^2} A_i \tag{65}$$

be extremal, subject to the constraint on the norm $\langle \psi | \psi \rangle = \int |\psi(A)|^2 dA = 1$.

The kinetic energy term is just the Laplacian with respect to a Euclidean metric on the space $\mathcal{A}$ of Hermitian matrices:

$$\tilde{g}(d[A_i]^a_b, d[A_j]^c_d) = \tilde{g}^{\ a\ c}_{ib\ jd} = K_{ij} \delta^a_d \delta^c_b. \tag{66}$$

This metric is invariant under the action of $U(N)$ on the matrix variables: $A_i \to U A_i U^{-1}$. Moreover, the potential energy $\operatorname{tr} V^I A_I$ is invariant under $U(N)$. We will regard this as a "gauge transformation," and allow only for wave functions that are also invariant under $U(N)$:

$$\psi(UAU^\dagger) = \psi(A). \tag{67}$$

The problem is to determine the eigenvalues and eigenfunctions of the Hamiltonian H over this class of invariant wave functions. Feynman diagram methods applied to this problem[19] suggest that there is an important simplification in the limit $N \to \infty$ holding V^I and K_{ij} fixed: only Feynman diagrams of planar topology contribute. This in turn implies that the gauge invariant observables such as $G_I = \frac{1}{N} \operatorname{tr} A_I$ have fluctuations of the order $\frac{1}{N^2}$: thus matrix models reformulated in terms of invariants tend to some classical theory in the large N limit. What are the configuration space, Poisson brackets and Hamiltonian of this classical theory?

A complete set of invariants for the matrices are given by $G_I = \frac{1}{N} \operatorname{tr} A_I$. For a fixed finite N, only a finite subset of the G_I can be independent variables. If we allow N to be arbitrary (may be even infinite) the G_I are independent except for cyclic symmetry, the reality conditions and inequalities. Thus the configuration space of matrix models is nothing but the space of noncommutative probability distributions that are tracial:

$$G_\phi = 1, \quad G^*_I = G_{\bar{I}}, \quad G_{IJ} f^{\bar{I}^*} f^J \geq 0 \text{ for any polynomial } f(X). \tag{68}$$

In essence our task is to change variables from Cartesian coordinates A^a_{ib} to "radial coordinates" G_I.

It is easier to change variables in the quadratic form $\langle \psi | H | \psi \rangle$ rather than the Hamiltonian operator H directly: we only have to deal with first order rathar than second order derivatives. By the chain rule of differentiation,

$$K_{ij} \frac{\partial \psi^*}{\partial A^a_{ib}} \frac{\partial \psi}{\partial A^b_{ja}} = K_{ij} \frac{\partial G_I}{\partial A^a_{ib}} \frac{\partial G_J}{\partial A^b_{ja}} \frac{\partial \psi^*}{\partial G_I} \frac{\partial \psi}{\partial G_J}. \tag{69}$$

Now,

$$\frac{\partial G_I}{\partial [A_i]_b^a} = \delta_I^{I_1 i I_2} [A_{I_2 I_1}]_a^b \,,$$

$$\Rightarrow \frac{\partial G_I}{\partial [A_i]_b^a} \frac{\partial G_J}{\partial [A_j]_a^b} = \delta_I^{I_1 i I_2} \delta_J^{J_1 j J_2} G_{I_2 I_1 J_2 J_1} \tag{70}$$

so that

$$K_{ij} \frac{\partial \psi^*}{\partial A_{ib}^a} \frac{\partial \psi}{\partial A_{ja}^b} = g_{I;J} \frac{\partial \psi^*}{\partial G_I} \frac{\partial \psi}{\partial G_J} \,, \tag{71}$$

where

$$g_{I;J} = K_{ij} \delta_I^{I_1 i I_2} \delta_J^{J_1 j J_2} G_{I_2 I_1 J_2 J_1} \,. \tag{72}$$

The geometrical meaning of this clear: $g_{I,J}$ is the inverse of the metric tensor induced on the space of noncommutative probability distributions by the Euclidean metric on the space of matrices.

To complete the story we need to understand as well the Jacobian of the transformation $A_i \to G_I$. This problem was studied extensively in a previous paper, where we showed that this Jacobian is $e^{N^2 \chi(G)}$ where $\chi(G)$ is the free entropy (in the sense of Voiculescu) of the noncommutative probability distribution whose moments are G_I. Thus we get

$$\langle \psi | H | \psi \rangle = \int \left[g_{I;J} \frac{\partial \psi^*}{N^2 \partial G_I} \frac{\partial \psi}{N^2 \partial G_J} + V^I G_I \psi^* \psi \right] e^{N^2 \chi(G)} dG \tag{73}$$

and, of course also,

$$\langle \psi | \psi \rangle = \int \psi^* \psi e^{N^2 \chi(G)} dG \,. \tag{74}$$

To take the limit as $N \to \infty$, it is useful to introduce a new "radial wave function"

$$\Psi = e^{\frac{1}{2} N^2 \chi} \psi \tag{75}$$

so that

$$\langle \psi | \psi \rangle = \int \Psi^* \Psi dG \,, \quad \int V^I G_I \psi^* \psi e^{N^2 \chi(G)} dG = \int V^I G_I \Psi^* \Psi dG \,. \tag{76}$$

In making this change of variables on the expectation value of kinetic energy, there will be some extra contributions from the derivatives of the entropy. These will add an extra term V_{coll} to the potential. It is our next task to determine this term. Using

$$\frac{\partial \psi}{N^2 \partial G_I} e^{\frac{1}{2} N^2 \chi} = \frac{\partial \Psi}{N^2 \partial G_I} - \frac{1}{2} \frac{\partial \chi}{\partial G_I} \Psi \,, \tag{77}$$

we get

$$g_{I;J} \frac{\partial \psi^*}{N^2 \partial G_I} \frac{\partial \psi}{N^2 \partial G_J} e^{N^2 \chi(G)} = g_{I;J} \frac{1}{N^2} \frac{\partial \Psi^*}{\partial G_I} \frac{1}{N^2} \frac{\partial \Psi}{\partial G_J} + \frac{1}{4} g_{I;J} \frac{\partial \chi}{\partial G_I} \frac{\partial \chi}{\partial G_J} |\Psi|^2$$

$$- \frac{1}{2} g_{I;J} \frac{\partial \chi}{\partial G_I} \frac{\partial |\Psi|^2}{N^2 \partial G_J} \,. \tag{78}$$

An integration by parts gives

$$g_{I;J} \frac{\partial \psi^*}{N^2 \partial G_I} \frac{\partial \psi}{N^2 \partial G_J} e^{N^2 \chi(G)} = g_{I;J} \frac{1}{N^2} \frac{\partial \Psi^*}{\partial G_I} \frac{1}{N^2} \frac{\partial \Psi}{\partial G_J} + \frac{1}{4} g_{I;J} \frac{\partial \chi}{\partial G_I} \frac{\partial \chi}{\partial G_J} |\Psi|^2$$

$$+ \frac{1}{2} \frac{1}{N^2} \frac{\partial}{\partial G_J} \left[g_{I;J} \frac{\partial \chi}{\partial G_I} \right] |\Psi|^2$$

$$+ \text{total divergence} \,. \tag{79}$$

The first term leads directly to the "radial" kinetic energy of Eq. (57). The last two terms do not involve derivatives of Ψ and hence are contributions to the potential arising from the change of variables: they represent a sort of "centrifugal potential" (or "collective potential" in the language of Jevicki–Sakita):

$$V_{\text{coll}} = \frac{1}{8} g_{I;J} \frac{\partial \chi}{\partial G_I} \frac{\partial \chi}{\partial G_J} + \frac{1}{4} \frac{1}{N^2} \frac{\partial}{\partial G_J} \left[g_{I;J} \frac{\partial \chi}{\partial G_I} \right] \,. \tag{80}$$

In the large N limit the second term will be small compared to the first. In any case, we show that it vanishes. Using the formula for $g_{I;J}$ and $\frac{\partial \chi}{\partial G_{Jj}} = G^{J;L} \eta_L^j$,

$$\frac{1}{4N^2} \frac{\partial}{\partial G_J} \left(g_{I;J} \frac{\partial \chi}{\partial G_I} \right) = \frac{1}{4N^2} K_{ij} \frac{\partial}{\partial G_I} \left(\delta_I^{I_1 i I_2} \delta_J^{J_1 j J_2} G_{I_2 I_1 J_2 J_1} \frac{\partial \chi}{\partial G_J} \right)$$

$$= \frac{1}{4N^2} K_{ij} \frac{\partial}{\partial G_{Ii}} \left(G_{IJ} \frac{\partial \chi}{\partial G_{Jj}} \right) = \frac{1}{4N^2} K_{ij} \frac{\partial \eta_I^j}{\partial G_{Ii}}$$

$$= \frac{1}{4N^2} K_{ij} \frac{\partial}{\partial G_{Ii}} (\delta_I^{I_1 j I_2} G_{I_1} G_{I_2}) = 0 \,, \tag{81}$$

since we are always differentiating with respect to a higher moment. Thus only the first term contributes to V_{coll}

$$V_{\text{coll}} = \frac{1}{8} K_{ij} G_{I_2 I_1 J_2 J_1} \frac{\partial \chi}{\partial G_{I_1 i I_2}} \frac{\partial \chi}{\partial G_{J_1 j J_2}}$$

$$= \frac{1}{8} K_{ij} G_{I_2 I_1 J_2 J_1} G^{I_2 I_1; L} \eta_L^i G^{J_2 J_1; M} \eta_M^j$$

$$= \frac{1}{8} K_{ij} G^{L;M} \eta_L^i \eta_M^j \,. \tag{82}$$

From our earlier discussion of the Fisher information we see that

$$V_{\text{coll}} = \frac{1}{4} K_{ij} \Phi^{ij} \,. \tag{83}$$

Now we can collect the different terms together to get the answer for the Hamiltonian stated in Eq. (57).

4. Approximate Large *N* Ground State and the Variational Principle

Our aim is to understand the ground state of the Hamiltonian

$$H = K_r + V_{\text{eff}} \,, \tag{84}$$

where $K_r = \frac{1}{2}K_{ij}\pi^{I_1 i I_2}G_{I_2 I_1 J_2 J_1}\pi^{J_1 j J_2}$ and $V_{\text{eff}} = \frac{1}{4}K_{ij}\Phi^{ij} + V^I G_I$. The large N ground state is the probability distribution minimizing the effective potential $V_{\text{eff}} = V_{\text{coll}} + V^I G_I$. In essence, this is the problem of minimizing free Fisher information holding the moments in the potential fixed via the Lagrange multipliers V^I. So far we have not made any approximations aside from the large N limit.

Mean field theory

The equation for minimization of the effective potential also defines for us a variational principle. If we minimize V_{eff} over a sub-manifold of the configuration space which corresponds to the choice of a few parameter family of actions for the static matrix model, we will get an upper bound for the ground state energy. But as we well know, explicit solutions of the Dyson–Schwinger equations for arbitrary multi-matrix models are not always possible. However we may chose as the action S some solvable model, which can then act as a variational ansatz. The simplest choice for the action is the Gaussian, which leads to the Wignerian sub-manifold of moments. This leads to an approximation method similar to mean field theory (MFT). As a first check on our methods, we present below mean-field calculations for some dynamical matrix models.

Since the Gaussian action for static matrix models leads to Wigner moments,

$$\Phi^{ij}(G) = \frac{1}{2}S^{iI}(G)S^{jJ}(G)G_{IJ} = \frac{1}{2}(G^{-1})^{ij} \tag{85}$$

on the Wignerian sub-manifold. Here $(G^{-1})^{ij}G_{jk} = \delta_k^i$. Let us specialize to matrix models with a quartic potential $V^{ij}G_{ij} + V^{ijkl}G_{ijkl}$. Then the effective potential to be minimized is

$$V_{\text{eff}}(G_{ij}) = \frac{1}{8}K_{ij}(G^{-1})^{ij} + V^{ij}G_{ij} + V^{ijkl}(G_{ij}G_{kl} + G_{il}G_{jk})\,. \tag{86}$$

The minimum of the effective potential is given by

$$\frac{\partial V_{\text{eff}}}{\partial G_{pq}} = -\frac{1}{8}K_{ij}(G^{-1})^{ip}(G^{-1})^{jq} + V^{pq}$$

$$+ V^{pqkl}G_{kl} + V^{ijpq}G_{ij} + V^{pjkq}G_{jk} + V^{ipql}G_{il} = 0\,. \tag{87}$$

Notice that this is a cubically nonlinear equation for the positive symmetric variational matrix G_{ij}. G_{ij} plays the role of a mean field and this is its self consistent equation.

4.1. *Mean field theory for dynamical one-matrix model*

To calibrate our approximation method, let us consider the dynamical one-matrix model

$$H = -\frac{1}{2}\frac{\partial^2}{\partial A_b^a \partial A_a^b} + \operatorname{tr}\left[\frac{1}{2}A^2 + gA^4\right], \tag{88}$$

whose exact large N ground state energy is known from the work of Brezin *et al.*[24] On the sub-manifold of Wigner semicircular distributions,

$$V_{\text{eff}}^{\text{sc}}(G_2) = \frac{1}{8G_2} + \frac{1}{2}G_2 + 2gG_2^2, \tag{89}$$

whose minimum occurs at

$$G_2^{\text{sc}}(g) = \frac{1}{24g}[-1 + \mu^{-1/3} + \mu^{1/3}], \tag{90}$$

where

$$\mu = -1 + 216g^2 + 12g\sqrt{-3 + 324g^2}. \tag{90}$$

The semicircular estimate is a variational upper bound $E_{\text{gs}} \leq E_{\text{gs}}^{\text{sc}}$.

4.2. *Cramér–Rao inequality*

Somewhat surprisingly, we can also get a lower bound for E_{gs}. For this we make use of the (noncommutative) Cramér–Rao (CR) inequality[22,23,12] along with the moment inequalities, which are conditions for positivity of the Hankel matrix. The CR inequality states that $\Phi^{ij} - (G^{-1})^{ij}$ is a positive matrix. It is the statement that among all distributions with fixed second moment, the Wigner distribution minimizes information. For one variable, it is the condition $\Phi G_2 \geq 1$ which is understood as a sort of uncertainty principle in statistics. The moment inequalities (MI) for a single variable are

$$G_2 \geq 0, \quad G_4 \geq G_2^2, \quad G_2G_4G_6 - G_4^3 + G_2^3G_6 - G_2^2G_4^2 \geq 0, \dots. \tag{91}$$

We can saturate the MI by choosing $G_2 \geq 0$; $G_4 = G_2^2$; $G_6 = G_2^3 \dots$. The MI are saturated by $\rho(x) = \delta(x - x_0)$, $G_{2n} = x_0^{2n}$, *not* the Wigner moments.

The CR-MI estimate is a lower bound $E_{\text{gs}}^{\text{CR-MI}} \leq E_{\text{gs}}$. We see from the following comparison with Brezin *et al.*'s exact $(N \to \infty)$ result, that in addition to providing bounds on E_{gs}, mean field theory is a reasonably good approximation.

g	$E_{\text{gs}}^{\text{CR-MI}}$	$E_{\text{gs}}^{\text{exact}}$	$E_{\text{gs}}^{\text{sc}}$
.01	.503	.505	.505
.1	.523	.542	.543
.5	.591	.651	.653
1	.653	.740	.743
50	1.794	2.217	2.236
1000	4.745	5.915	5.969
$g \to \infty$	.472 $g^{1/3}$	.590 $g^{1/3}$	.595 $g^{1/3}$

4.3. *Semicircular ansatz for a dynamical two-matrix model*

We consider the two matrix model, that is of considerable interest in M theory[1] (for simplicity we look at the nonsupersymmetric counterpart of the BFSS model), defined by

$$H = -\frac{1}{2}\left(\frac{\partial^2}{\partial A_a^b \partial A_b^a} + \frac{2}{\partial B_a^b \partial B_b^a}\right) + V(A, B),$$

where

$$V(A, B) = \text{tr}\left[\frac{m^2}{2}(A^2 + B^2) + \frac{c}{2}(AB + BA) - \frac{g}{4}[A, B]^2\right]. \qquad (92)$$

We read off $V^{ij} = \begin{pmatrix} \frac{1}{2}m^2 & \frac{1}{2}c \\ \frac{1}{2}c & \frac{1}{2}m^2 \end{pmatrix}$, $V^{ABAB} = -V^{ABBA} = -\frac{1}{2}g$ and consider the region $g \geq 0$, $|c| \leq m^2$ where V^{ij} is positive. The effective potential to be minimized is

$$V_{\text{eff}} = \frac{1}{8}\text{tr}[G^{-1}] + \frac{1}{2}m^2(G_{AA} + G_{BB}) + cG_{AB} - \frac{1}{2}g(G_{ABAB} - G_{ABBA}). \qquad (93)$$

In terms of the positive $G_{ij} = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$ ($|\beta| \leq \alpha$) parametrizing the Wignerian sub-manifold of configuration space,

$$V_{\text{eff}}(\alpha, \beta) = \frac{\alpha}{4(\alpha^2 - \beta^2)} + m^2\alpha + c\beta + \frac{1}{2}g(\alpha^2 - \beta^2). \qquad (94)$$

For the above range of parameters, V_{eff} is bounded below. Since the information and potential scale are opposite, it has a minimum at a nontrivial $\alpha > 0$. For $c = 0$, we get $\beta = 0$ and a cubic for α : $4g\alpha^3 + 4m^2\alpha^2 - 1 = 0$. For $c > 0$ the actual solution of the algebraic equations can be obtained numerically.

4.4. *MFT for relativistic $\lambda\phi^4$ matrix field theory*

Next we consider a relativistic Hermitian scalar matrix field $\phi_b^a(x)$ in $d + 1$ dimensions. The Hamiltonian in the Schrödinger representation is

$$H = \int d^dx \left[-\frac{1}{2}\frac{\delta^2}{\delta\phi_b^a(x)\delta\phi_a^b(x)} + \text{tr}\left(\frac{1}{2}|\boldsymbol{\nabla}\phi(x)|^2 + \frac{m_0^2}{2}\phi(x)^2 + \lambda_0\phi(x)^4\right)\right]. \qquad (95)$$

Let $G_{xy} = \frac{\text{tr}}{N}\phi(x)\phi(y)$, $\int d^d y\, G^{xy}G_{yz} = \delta^d(x-z)$. Within the mean field approximation, the effective potential is

$$V_{\text{eff}} = \frac{1}{8}\int dx\, G^{xx} + \int dx\, dy \frac{1}{2}\delta(x-y)\left(-\frac{\partial^2}{\partial x^2} + m_0^2\right)G_{xy} + 2\lambda_0 \int dx (G_{xx})^2\,. \tag{96}$$

Minimizing it leads to an integral equation for the mean field G_{xy}

$$-\frac{1}{8}\delta(x-y) + \frac{1}{2}\int dz\, G_{zx}\left(-\frac{\partial^2}{\partial z^2} + m_0^2\right)G_{zy} + 4\lambda_0 \int dz\, G_{zz}G_{zx}G_{zy} = 0\,. \tag{97}$$

Assuming a translation invariant ground state

$$G_{xy} = G(x-y) = \int \frac{d^d p}{(2\pi)^d} e^{ip(x-y)}\tilde{G}(p)\,, \tag{98}$$

this equation can be solved:

$$\tilde{G}(p) = \frac{1}{2\sqrt{p^2 + m^2}}\,, \tag{99}$$

where m is determined self consistently by

$$m^2 = m_0^2 + 4\lambda_0 \int^\Lambda \frac{d^d p}{(2\pi)^d}\frac{1}{\sqrt{p^2 + m^2}}\,. \tag{100}$$

We recognize this as the mass renormalization; log, linearly and quadratically divergent in one, two and three spatial dimensions. We need to go beyond the mean field ansatz to see coupling constant renormalization. We will address these issues in a separate paper.

Acknowledgment

References

1. T. Banks, W. Fishler, S. H. Shenker and L. Susskind, *Phys. Rev.* **D55**, 5112 (1997).
2. A. Jevicki and S. Das, *Mod. Phys. Lett.* **A5**, 1639 (1990).
3. A. Jevicki and J. Rodrigues, *Nucl. Phys.* **B421**, 278 (1994).
4. D. V. Voiculescu, K. J. Dykema and A. Nica, *Free Random Variables* (American Mathematical Society, Providence, USA, 1992).
5. D. Gross and R. Gopakumar, *Nucl. Phys.* **B451**, 379 (1995).
6. M. Douglas, *Nucl. Phys. Proc. Suppl.* **41**, 66 (1995).
7. M. Douglas, *Phys. Lett.* **B344**, 117 (1995).
8. L. Akant, G. S. Krishnaswami and S. G. Rajeev, *Int. J. Mod. Phys.* **A17**, 2413 (2002).
9. D. Voiculescu, *Invent. Math.* **118**, 411 (1994).
10. D. Voiculescu, *Invent. Math.* **132**, 189 (1998).
11. D. Voiculescu, *Indiana U Math. J.* **49**, 837 (2000).
12. D. Voiculescu, *Lecture Notes in Math.*, Vol. 1738 (Springer, Berlin, 2000).

13. B. Sakita, *Phys. Rev.* **D21**, 1067 (1980).
14. A. Jevicki and B. Sakita, *Nucl. Phys.* **B165**, 511 (1980).
15. A. Jevicki and B. Sakita, *Phys. Rev.* **D22**, 467 (1980).
16. A. Jevicki and B. Sakita, *Nucl. Phys.* **B185**, 89 (1981).
17. B. Sakita, *Quantum Theory of Many Variable Systems and Fields* (World Scientific, Singapore, 1985).
18. L. G. Yaffe, *Rev. Mod. Phys.* **54**, 407 (1982).
19. G. 't Hooft, *Nucl. Phys.* **B72**, 461 (1974).
20. A. Agarwal and S. G. Rajeev, *Mod. Phys. Lett.* **A17**, 481 (2002).
21. M. L. Mehta, *Random Matrices* (Academic, New York, 1967).
22. C. R. Rao, *Linear Statistical Inference and its Applications* (Wiley, New York, 1965).
23. E. Pitman, *Some Basic Theory for Statistical Inference* (Chapman and Hall, London, 1979).
24. Brezin *et al.*, *Commun. Math. Phys.* **59**, 35 (1978).