ELSEVIER

Nuclear Physics B 474 [FS] (1996) 726–744

NUCLEAR PHYSICS B [FS]

Law of addition in random matrix theory

A. Zee

Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106-4030, USA

Received 27 February 1996; revised 14 May 1996; accepted 21 May 1996

Abstract

We discuss the problem of adding random matrices, which enables us to study Hamiltonians consisting of a deterministic term plus a random term. Using a diagrammatic approach and introducing the concept of "gluon connectedness", we calculate the density of energy levels for a wide class of probability distributions governing the random term, thus generalizing a result obtained recently by Brézin, Hikami and Zee. The method used here may be applied to a broad class of problems involving random matrices.

PACS: 02.50; 5.20; 11.10; 71.20
Keywords: Random matrices; Addition; Deterministic plus random

1. Introduction

Some four decades ago, Wigner [1] proposed studying the distribution of energy levels of a random Hamiltonian given by

$$H = \varphi , \tag{1}$$

where φ is an N by N hermitian matrix taken from the distribution

$$P(\varphi) = \frac{1}{Z} e^{-N \mathrm{tr}\, V(\varphi)} , \tag{2}$$

with Z fixed by $\int d\varphi P(\varphi) = 1$. This problem has been studied intensively by Dyson, Mehta, and others over the years [2–4]. Two years ago, Brézin and Zee discovered that, remarkably, while the density of eigenvalues depends [5] on V, the correlation between the density of eigenvalues, when suitably scaled, is independent [6] of V. This universality was also obtained earlier and independently by Ambjørn and collaborators [7]. Since then, it has been clarified and extended by other authors [8,24,10], studied numerically [11], and furthermore, shown to hold even when the distribution (2) is

PII S0550-3213(96)00276-3

generalized to a much broader class of distributions [12]. We expect that the discussion to be given below will hold also for this broader class of distributions, but for the sake of simplicity we will not work this through here.

In recent work [13], Brézin and Zee have generalized this Wigner problem to the case of a Hamiltonian given by the sum of a deterministic term and a random term

$$H = H_0 + \varphi . \tag{3}$$

Here H_0 is a diagonal matrix with diagonal elements ϵ_i, $i = 1, 2, \ldots, N$, and φ a random matrix taken from the ensemble (2). For the Gaussian case, namely with $V(\varphi) = \frac{1}{2}\varphi^2$, Pastur [26] has long ago determined the density of eigenvalues. The work described in [13] went beyond Pastur's work in that the correlation function between the density of eigenvalues in the Gaussian case was also determined. More recently, in a work with Brézin and Hikami [14], we managed to determine the density of eigenvalues for $V(\varphi) = \frac{1}{2}\varphi^2 + g\varphi^4$ to all orders in g. The correlation function was also computed, but only to first order in g.

This problem of "determinism plus chance" may be regarded as a generic problem in physics, and as such represents a significant generalization of Wigner's problem. For example, consider an electron moving in a magnetic field and scattering off impurities. We note that these "deterministic plus random" problems are considerably more difficult than the purely random problems defined in (1) and (2). A standard approach to solving the purely random problem involves diagonalizing the random matrix φ and then use orthogonal polynomials to disentangle the resulting expression. Clearly, in (3) we cannot diagonalize φ without un-diagonalizing H_0 and thus the orthogonal polynomial approach fails.

In this paper, we point out that the problem given in (3) is a special case of a broader class of problems involving the addition of random matrices. The deterministic Hamiltonian H_0 may in turn be replaced by a random Hamiltonian. Indeed, a deterministic matrix is but a special case of a random matrix. We will extend the work of Brézin, Hikami, and Zee [14] and determine the density of eigenvalues for the Hamiltonian given in (3) for an arbitrary V.

Our work is inspired by recent advances in the mathematical literature involving the theory of non-commutative probability and operator algebra [15–17]. A number of physicists have already brought these advances to the attention of the physics community [18,19,22]. While our work is thus inspired, we will not be using the mathematical approach given in [15–17], but instead will be based on the diagrammatic approach developed in [13] and subsequent work [20,14].

2. Adding random matrices

Consider a Hamiltonian given by

$$H = \varphi_1 + \varphi_2 \tag{4}$$

with the matrices $\varphi_{1,2}$ taken from the probability distribution

$$P(\varphi_1,\varphi_2) = \frac{1}{Z} e^{-N\mathrm{tr}[V_1(\varphi_1)+V_2(\varphi_2)]} \equiv P_1(\varphi_1)P_2(\varphi_2). \tag{5}$$

Notice that the probability distribution factorizes. (This is known as "free" in the mathematical literature.) The problem defined in (3) represents a special case. Previously, with D'Anna and with Brézin we have studied the problem given in (4) but for the more difficult case [21,9] in which $P(\varphi_1,\varphi_2)$ contains terms linking φ_1 and φ_2. Indeed, a detailed determination of the correlation function over all "distance scales" is a non-trivial problem even for Gaussian distributions [9]. The discussion in this paper goes through precisely because φ_1 and φ_2 do not couple to each other in $P(\varphi_1,\varphi_2)$.

Let us now mention a few necessary definitions. Define the Green's function

$$G(z) \equiv \left\langle \frac{1}{N}\mathrm{tr}\frac{1}{z-H} \right\rangle = \int\int d\varphi_1 d\varphi_1 P(\varphi_1,\varphi_2)\frac{1}{N}\mathrm{tr}\frac{1}{z-(\varphi_1+\varphi_2)}. \tag{6}$$

The density of eigenvalues is then given by $\rho(\mu) = \left\langle \frac{1}{N}\mathrm{tr}\delta(\mu - H)\right\rangle = -\frac{1}{\pi}\mathrm{Im}G(\mu+i\epsilon)$. In this paper we focus on the density of eigenvalues, leaving the correlation for a future work. Note that the factors of N are chosen in our definitions such that the interval over which $\rho(\mu)$ is non-zero is finite (i.e. of order N^0) in the large N limit.

We may regard the distribution (5) as defining a $(0+0)$-dimensional field theory. The Feynman diagram expansion is then simply obtained by expanding $G(z)$ in inverse powers of z and doing the integrals in (6). As explained in [13], it is useful to borrow the terminology of large N quantum chromodynamics [25] from the particle physics literature, and speak of quark and gluon lines. See Fig. 1 for a graphical representation. (It is of course not necessary to use this language, and readers not familiar with this language can simply think of the diagrams as representing the different terms one encounters in doing the integral in (6).) The quark propagator simply comes from the explicit factor of z in (6) and is represented by a single line and given by $1/z$. The quadratic terms in $V_1(\varphi_1)$ and $V_2(\varphi_2)$ determine the gluon propagator, represented by double lines. Here, in a minor departure from large N quantum chromodynamics we have two types of gluons, corresponding to φ_1 and φ_2. The gluon propagators are proportional to

$$\left\langle \varphi^i_{\alpha j}\varphi^k_{\beta l}\right\rangle \propto \delta_{\alpha\beta}\delta^i_l\delta^k_j\frac{1}{N}. \tag{7}$$

The non-Gaussian terms in $V_1(\varphi_1)$ and $V_2(\varphi_2)$ describe the interaction between the gluons.

The reason that we can solve this problem is because, while the two types of gluons have arbitrarily complicated interactions among themselves, they do not interact with each other. Note that while the gluons both interact with the quark, our problem is such that we do not have to include quark loops and thus the quark does not induce interaction between the two gluons. This is clear from the definition of our problem. Another way of saying this is to note that the Green's function may be represented, by using the replica trick, as

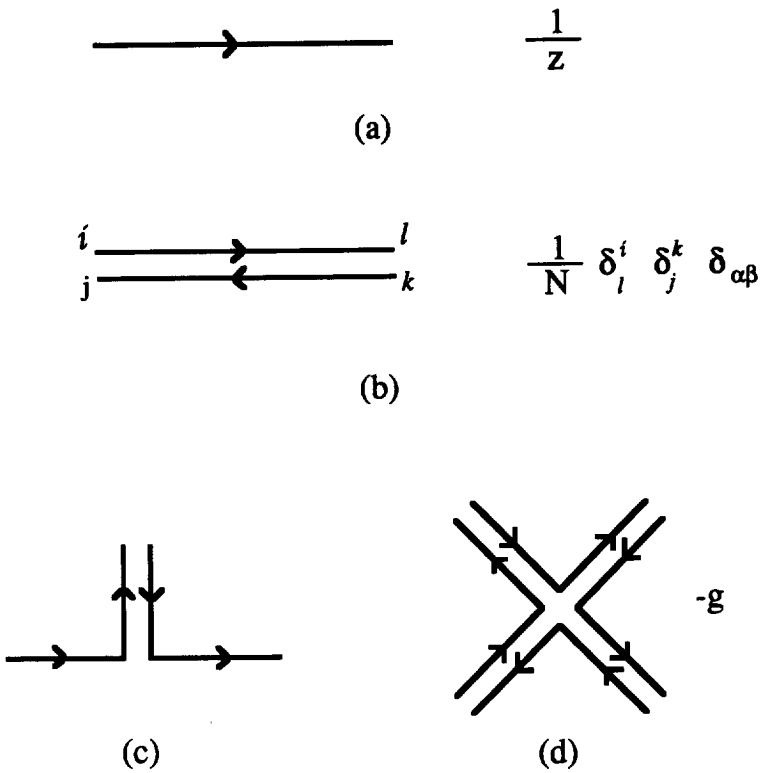

Fig. 1. Feynman rules: (a) quark propagator, (b) gluon propagator, (c) quark gluon vertex, (d) gluon interaction, illustrated here with a $g\varphi^4$ vertex.

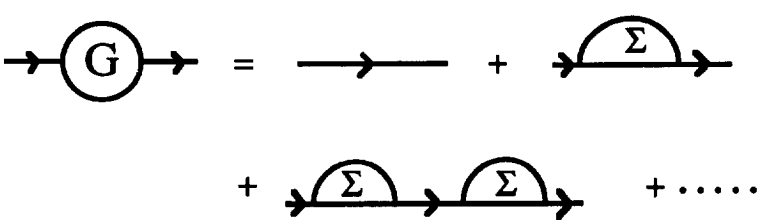

Fig. 2. Quark propagator and one-particle irreducible self-energy.

$$G(z) = \lim_{n\to 0} \int D\psi^\dagger D\psi D\varphi P(\varphi)\psi_1^\dagger\psi_1 \exp\left[-\sum_{\alpha=1}^{n}\psi_\alpha^\dagger(z-\varphi)\psi_\alpha\right]. \tag{8}$$

Note that in this language the ψ's represent the quark fields and φ the gluon fields. The interaction between gluon and quarks are given by $\psi_\alpha^\dagger\varphi\psi_\alpha$. (Color indices are suppressed here.) The interaction of the gluons with each other is determined by $P(\varphi)$. Since internal quark loops are proportional to the number of replicas n, they vanish in the $n \to 0$ limit.

Let us then calculate the Green's function, which as usual can be written as (see Fig. 2)

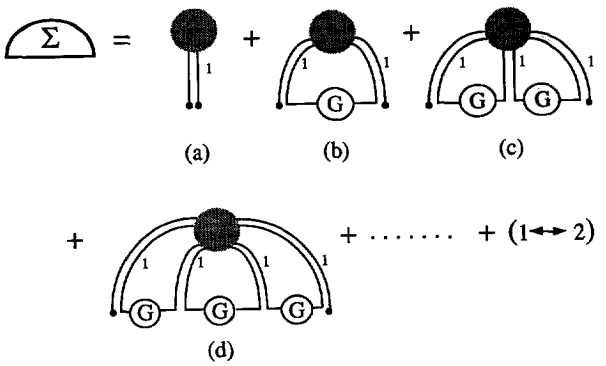

Fig. 3. Quark self-energy: the gluons shown explicitly are all of type 1. There are of course also type 2 gluons inside the quark propagator G.

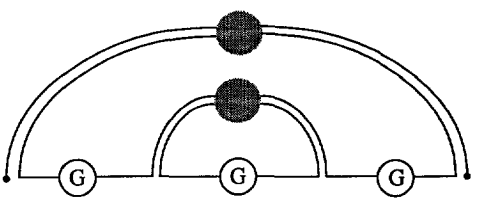

Fig. 4. A class of diagrams not included in (3d).

$$G(z) = \frac{1}{z - \Sigma(z)} \tag{9}$$

in terms of the one-particle irreducible self-energy $\Sigma^i_j(z) = \delta^i_j \Sigma(z)$. The self-energy is then determined by the set of diagrams in Fig. 3 with the corresponding equation

$$\begin{aligned}\Sigma(z) &= \langle \frac{1}{N} \mathrm{tr}\, \varphi_1 \rangle_{\mathrm{gc}} + \langle \frac{1}{N} \mathrm{tr}\, \varphi_1^2 \rangle_{\mathrm{gc}} G(z) + \langle \frac{1}{N} \mathrm{tr}\, \varphi_1^3 \rangle_{\mathrm{gc}} G(z)^2 + \ldots + (1 \leftrightarrow 2) \\ &= \sum_{k=1}^{\infty} \langle \frac{1}{N} \mathrm{tr}\, \varphi_1^k \rangle_{\mathrm{gc}} G(z)^{k-1} + (1 \leftrightarrow 2) \\ &= \frac{1}{G} \left[\langle \frac{1}{N} \mathrm{tr} \frac{1}{1 - \varphi_1 G} \rangle_{\mathrm{gc}} - 1 \right] + (1 \leftrightarrow 2) \\ &= \frac{1}{G} \left[\frac{1}{G} G_{\mathrm{gc}1}(\frac{1}{G}) - 1 \right] + (1 \leftrightarrow 2). \end{aligned} \tag{10}$$

In order to write this equation, we have to invoke the factorization of $P(\varphi_1, \varphi_2)$, which tells us that the two kinds of gluons do not interact, and the large N limit, which tells us that the two kinds of gluon lines cannot cross.

We are led to introduce in (10) the notion of "gluon connectedness", denoted by "gc" henceforth. The necessity for this notion is illustrated by the shaded blob describing the interaction of the gluons in Fig. 3d: it should not include the diagram shown in Fig. 4; this class of diagrams is already included in Fig. 3b. In other words, a gluon connected blob with k external gluon lines is such that it cannot be separated into two blobs, with

k_1 gluon lines and k_2 gluon lines respectively (with $k_1 + k_2 = k$ of course). In the last line we have defined the "gluon connected Green's function"

$$G_{gc1}(z) = \langle \frac{1}{N} \mathrm{tr} \frac{1}{z - \varphi_1} \rangle_{gc} \tag{11}$$

and similarly $G_{gc2}(z)$. The operations implied in (11) are clearly allowed since $\langle \frac{1}{N} \mathrm{tr} \cdot \rangle_{gc}$ is a linear operation. Note also that we have not assumed that V_α is an even function of its argument. In particular, we include a possible tadpole term indicated by $\langle \frac{1}{N} \mathrm{tr} \varphi_\alpha \rangle_{gc}$ in (10).

We should emphasize that the shaded blobs include interactions between gluons to all orders. It is very complicated, if not hopeless, to calculate these blobs in terms of V_1 and V_2, but fortunately, as we will show below, we do not have to calculate them explicitly. In our previous papers, we regarded the cubic, quartic, and so on, terms in V_α as interactions and proceeded to calculate the Green's function and correlation function in terms of the various coupling constants. We follow a different strategy here, and try to express the Green's function $G(z)$ directly in terms of $G_1(z)$ and $G_2(z)$ where

$$G_\alpha(z) \equiv \langle \frac{1}{N} \mathrm{tr} \frac{1}{z - \varphi_\alpha} \rangle \tag{12}$$

for $\alpha = 1, 2$ are the Green's functions for two separately and purely random problems. (The average in (12) is performed with the distribution $P_\alpha(\varphi_\alpha) = (1/Z_\alpha) e^{-N \mathrm{tr} V_\alpha(\varphi_\alpha)}$ of course.) In this way, we attempt to bypass having to deal with V_1 and V_2 altogether.

To see how to do this, let us go back to the simpler problem defined by (1) and (2). Following the same diagrammatic analysis leading to (10) we find that the Green's function $G(z)$ and self-energy $\Sigma(z)$ for this simpler problem are related by

$$\Sigma(z) = \frac{1}{G} \left[\frac{1}{G} G_{gc}(\frac{1}{G}) - 1 \right] \tag{13}$$

with, evidently,

$$G_{gc}(z) \equiv \langle \frac{1}{N} \mathrm{tr} \frac{1}{z - \varphi} \rangle_{gc} . \tag{14}$$

Combining (13) and (9) we find

$$\frac{1}{G^2} G_{gc}(\frac{1}{G}) = z . \tag{15}$$

For the sake of convenience, we may, with due respect to Green, somewhat fancifully define a "Blue's function" by

$$B(z) \equiv \frac{1}{z^2} G_{gc}(\frac{1}{z}) . \tag{16}$$

Thus, we learn from (15) that the Blue's function is the functional inverse of the Green's function

$$B(G(z)) = z . \tag{17}$$

From the normalization of the probability distribution $P(\varphi)$ we obtain trivially the "sum rule" $G(z) \to 1/z$ as $z \to \infty$, thus implying that the Blue's function $B(z) \to 1/z$ as $z \to 0$.

Let us now go back to the more involved problem defined by (4). First, we define for $\alpha = 1, 2$ two Blue's functions B_α as the functional inverse of G_α, respectively. We now see that (10), when combined with (9), says simply that

$$z + \frac{1}{G} = B_1(G) + B_2(G) . \tag{18}$$

Thus, the law of addition for the Blue's function is given by

$$B_{1+2}(z) = B_1(z) + B_2(z) - \frac{1}{z} . \tag{19}$$

This equation tells us how to obtain the Blue's function associated with $\varphi_1 + \varphi_2$ from the Blue's functions associated with φ_1 and φ_2.

The procedure for determining the Green's function and hence the density of eigenvalues of the problem defined by (1) and (2) is then as follows: given the Green's functions G_1 and G_2, determine the corresponding Blue's functions B_1 and B_2 by functionally inverting G_1 and G_2 respectively, calculate B_{1+2} according to (19), then determine the functional inverse of B_{1+2} to find the desired Green's function $G(z)$.

Let us remark briefly on the connection to the mathematical literature. Voiculescu [15] has introduced the "R transform". It turns out that the R function discussed by mathematicians is simply related to B by $B(z) = 1/z + R(z)$. In fact, we see that the Dyson–Schwinger equation (9) when combined with (17) gives simple $B(G(z)) = \frac{1}{G(z)} + \Sigma(z)$. Thus, the R function of the mathematicians is nothing but the self-energy Σ of the physicists expressed in terms of different arguments:

$$R(G(z)) = \Sigma(z) . \tag{20}$$

3. Addition algorithm at work

Let us see how the addition algorithm works from a simple example. For $P(\varphi)$ Gaussian (that is, $V(\varphi) = \mathrm{tr}\,\frac{1}{2}\varphi^2$), it is well known (see for example Ref. [13]) that the Green's function is determined by

$$z = G(z) + \frac{1}{G(z)} . \tag{21}$$

In other words,

$$G(z) = \frac{1}{2}(z - \sqrt{z^2 - 4}) . \tag{22}$$

Substituting $z \to B$ into (21), we obtain immediately that

$$B(z) = z + \frac{1}{z} . \tag{23}$$

Thus, in this simple case, the Blue's function B, which of course contains the same information as the Green's function G, actually has a simpler form than G.

Using dimensional analysis we see immediately that for $V(\varphi) = \frac{1}{2}m^2\varphi^2$ the Blue's function reads

$$B(z) = \frac{z}{m^2} + \frac{1}{z}. \tag{24}$$

As a particularly simple example, consider adding two Gaussian random matrices φ_1 and φ_2 taken from distributions defined by $V_i(\varphi_i) = \frac{1}{2}m_i^2\varphi_i^2$, $i = 1, 2$. Then from (23) and from the addition law in (19) immediately follows the usual Gaussian law of addition, $m_{1+2}^{-2} = m_1^{-2} + m_2^{-2}$. Of course, to obtain this result, we hardly need the formalism given in this paper. This simple exercise provides a slight check on the formalism.

We will now show that the addition formalism given here applies to the addition of a deterministic matrix and a random matrix taken from a distribution of the form in (2), even though our proof was given for the addition of two random matrices. Let us first determine the Blue's function corresponding to non-random matrices. In the most trivial case, consider φ to be a constant c times the unit matrix. Then from the Green's function $G(z) = 1/(z - c)$ we find the Blue's function $B(z) = c + 1/z$. For a slightly less trivial example, let φ be a diagonal matrix with matrix elements given by ϵ_i with $i = 1, \ldots, N$. The Green's function is given by

$$G(z) = \frac{1}{N}\sum_i \frac{1}{z - \epsilon_i}. \tag{25}$$

Then the corresponding Blue's function is determined (implicitly) by

$$\frac{1}{N}\sum_i \frac{1}{B(z) - \epsilon_i} = z. \tag{26}$$

Now consider the problem defined in (3). Since we know from (26) and (23) the Blue's functions corresponding to the two terms in the Hamiltonian, we learn immediately from (19) the Blue's function for H:

$$B_{1+2}(z) = B_1(z) + z + \frac{1}{z} - \frac{1}{z} = B_1(z) + z, \tag{27}$$

with B_1 determined by (26) with the renaming substitution $B \to B_1$. The desired Green's function $G(z)$ is now determined by solving for the functional inverse of the function B_{1+2}, that is, by the equation

$$B_{1+2}(G) = z, \tag{28}$$

or equivalently, upon substituting $z \to G$ in (27), $z = B_1(G) + G$, which we write more conveniently for our next step as

$$B_1(G) = z - G. \tag{29}$$

Let us now evaluate the two sides of this equation with the function $G_1(\cdot)$. Since $G_1(B_1(G(z))) = G(z)$ we obtain immediately

$$G(z) = G_1(z - G(z)) . \tag{30}$$

Noting that G_1 is given by (25) with the substitution $G \to G_1$, we have

$$G(z) = \frac{1}{N} \sum_i \frac{1}{z - \epsilon_i - G(z)} , \tag{31}$$

precisely the classic result of Pastur [26] which was obtained diagrammatically in [13].

We can now go on and show that the addition formalism also applies to the general version of the problem defined in (3): find the density of energy levels of a Hamiltonian given by $H = H_0 + \varphi$ with φ drawn from the general distribution (2). With a slight shift in notation, let us call the Blue's function associated with H_0 and with φ respectively B_0 and B_2. Then the Blue's function associated with H is given by $B(z) = B_0(z) + B_2(z) - 1/z$. Substituting in this equation $z \to G(z)$ (where $G(z)$ is the unknown Green's function associated with H), we find immediately that

$$B_0(G) = z + \frac{1}{G} - B_2(G) . \tag{32}$$

Anticipating the next step, we define

$$\Sigma(z) \equiv B_2(G(z)) - \frac{1}{G(z)} . \tag{33}$$

With this definition, we write (32) as

$$B_0(G) = z - \Sigma(z) . \tag{34}$$

Let us now evaluate both sides of (32) with the Green's function $G_0(\cdot)$ associated with H_0. We find instantly that

$$G_0(B_0(G(z))) = G(z) = G_0(z - \Sigma(z)) = \frac{1}{N} \sum_i \frac{1}{z - \epsilon_i - \Sigma(z)} , \tag{35}$$

which we write as

$$G(z) = \frac{1}{N} \sum_i \frac{1}{z - \epsilon_i - \Sigma(z)} \tag{36}$$

for future reference.

Let us repeat this trick: rewrite (33) as $B_2(G(z)) = \Sigma(z) + 1/G(z)$ and evaluate both sides with the function $G_2(\cdot)$. We obtain

$$G(z) = G_2\left(\Sigma(z) + \frac{1}{G(z)}\right) . \tag{37}$$

These two equations, (36) and (37), allow us to determine the two unknown functions $G(z)$ and $\Sigma(z)$, provided we know the Green's function $G_2(z)$. But what is the Green's

function $G_2(z)$? It is just the Green's function associated with the random matrix φ drawn from the general distribution (2). But this was obtained by Brézin et al. [5] almost twenty years ago. These authors told us that (for $V(z)$ an even polynomial for the sake of notational simplicity)

$$G_2(z) = \frac{1}{2}[V'(z) - P(z)\sqrt{z^2 - a^2}]\,. \tag{38}$$

Here $V'(z) \equiv dV/dz$, $P(z)$ is a polynomial, and a determines the end points of the spectrum of eigenvalues. The quantities $P(z)$ and a are determined [1] by the "sum rule" that $G_2 \to 1/z$ as $z \to \infty$.

In summary, and repeating various equations for clarity, we have obtained the following result. For a Hamiltonian of the form $H = H_0 + \varphi$ with the random matrix φ drawn from an arbitrary distribution defined by $V(\varphi)$ (taken to be even for simplicity), we can determine the Green's function $G(z)$ and hence the density of eigenvalues by solving simultaneously the two equations

$$G(z) = \frac{1}{N}\sum_i \frac{1}{z - \epsilon_i - \Sigma(z)} \tag{42}$$

and

$$G(z) = G_2\left(\Sigma(z) + \frac{1}{G(z)}\right)\,, \tag{43}$$

where the function G_2 is given by (38). In general, for an arbitrary set of ϵ_i's and a non-Gaussian V, these equations can only be solved numerically.

It is clearly of some notational benefit to give the combination appearing in (43) a name: $\sigma(z) \equiv \Sigma(z) + 1/G(z)$. We can then simplify (43) slightly to [2]

$$P^2(\sigma)(\sigma^2 - a^2) = (V'(\sigma) - 2G)^2\,. \tag{44}$$

[1] For the sake of completeness, even though we will not need these explicit formulas in the text, let us record (and for the sake of simplicity, with V an even function) that for $V(\varphi) = \sum_{k=1}^{p} \frac{1}{2k} g_k \varphi^{2k}$ we have

$$P(z) = \frac{1}{2}\sum_{k=1}^{p} g_k \sum_{n=0}^{k-1} \frac{(2n)!}{(n!)^2}\left(\frac{a^2}{4}\right)^n z^{2k-2n-2}\,, \tag{39}$$

$$\frac{1}{2}\sum_{k=1}^{p} g_k \frac{(2k)!}{(k!)^2}\left(\frac{a^2}{4}\right)^k = 1\,. \tag{40}$$

In particular, for $V = \frac{1}{2}\varphi^2 + \frac{g}{4}\varphi^4$ we have

$$G(z) = \frac{1}{2}\left[z + gz^3 - \left(1 + \frac{a^2 g}{2} + gz^2\right)\sqrt{z^2 - a^2}\right] \tag{41}$$

and $a^2 = \frac{2}{3g}(\sqrt{1 + 12g} - 1) = 4(1 - 3g + 18g^2 + \ldots)$.

[2] It is perhaps worth noting that B_2 satisfies an equation similar to (44), namely $P^2(B_2)(B_2^2 - a^2) = (V'(B_2) - 2z)^2$.

Thus, we can use (44) to determine σ, and hence Σ, in terms of G. Plugging this into (42) then gives us an equation for G.

As mentioned earlier, Brézin, Hikami and Zee [14] recently used the equation of motion method and a detailed diagrammatic analysis to determine the Green's function for the problem in (3) with the distribution defined by $V(\varphi) = \frac{1}{2}\varphi^2 + g\varphi^4$. It is straightforward, although slightly tedious, to verify that for this simple case (44) reduces to Eq. (4.15) in [14]. The analysis given here is considerably simpler.

Thus, we have shown a posteriori that the addition algorithm given here also applies to the problem of adding a deterministic matrix and a random matrix taken from a distribution of the form in (2). (Incidentally, the algorithm obviously cannot be applied to the addition of two deterministic matrices since, given the density of eigenvalues for each matrix, we have no way of knowing the relative orientation between the two matrices.) To conclude this section, it is perhaps worthwhile, for the sake of completeness, to sketch how (36) was derived in [14] (Eq. (4.7) there.) Using the notation in (3) we define the matrix Green's function

$$G^i_j(z) \equiv \langle \left(\frac{1}{z-H}\right)^i \rangle_j , \qquad G^i_{0j}(z) \equiv \left(\frac{1}{z-H_0}\right)^i_j .$$

Eq. (9) is immediately generalized to its matrix version:

$$G(z) = \left((G_0(z))^{-1} - \Sigma(z)\right)^{-1} . \tag{45}$$

By inspecting the appropriate analog of Fig. 3, we see that even though $G(z)$ and $G_0(z)$ are not proportional to the unit matrix, the matrix Σ^i_j is still proportional to the unit matrix: $\Sigma^i_j(z) = \delta^i_j \Sigma(z)$. Taking the trace of (45), we obtain precisely (36). (Here and henceforth, $G(z)$ and $\Sigma(z)$ again denote scalar functions rather than matrix functions.) Note that (36) reduces to (31) in the Gaussian case.

4. Energy bands

The function B contains the same amount of information as G, since one function is the inverse of the other. As an example, consider the problem of determining the end points of the energy spectrum. Near the end point, call it a, the density of states $\rho(\mu)$, that is, the imaginary part of $G(z)$, vanishes generically like $(a-\mu)^{1/2}$. Thus, the end point is determined by the equation

$$\left.\frac{dG}{dz}\right|_{z=a} = \infty . \tag{46}$$

Now consider the defining equation for B: namely $G(B(w)) = w$. Differentiating, we obtain

$$\frac{dG}{dB}\frac{dB}{dw} = 1 . \tag{47}$$

Thus, we obtain an alternative equation for determining the end points a of the energy spectrum of a random Hamiltonian: solve the equation

$$\left.\frac{dB}{dw}\right|_{B=a} = 0 . \tag{48}$$

Let us now apply these simple considerations to determine the end points of the spectrum of a Hamiltonian of the form $H = H_0 + \varphi$.

As remarked above, with an arbitrary H_0 (that is, an arbitrary set of ϵ_i's) and a general V, we can hardly expect to solve (42) and (43) analytically. Let us retreat to the case in which φ is taken from a Gaussian distribution, so that $\Sigma(z)$ in (42) can be replaced by $G(z)$. (In other words, the solution of (43) is simply $\Sigma(z) = G(z)$.) Thus, we have to solve

$$G(z) = \frac{1}{N}\sum_i^N \frac{1}{z - \epsilon_i - G(z)} . \tag{49}$$

We will be interested in the case in which the Hamiltonian H_0 exhibits degeneracy. Let K be the number of distinct ϵ_i's (with $K \leqslant N$ of course.) To determine G we have to solve a polynomial equation of degree $K + 1$.

As explained above, in general, given an equation determining $G(z)$, we simply substitute $z \to B(w)$ and $G(z) \to w$ into that equation to obtain the equation for determining $B(w)$. Thus, in the present example, the function B is determined by

$$w = \frac{1}{N}\sum_i^N \frac{1}{B(w) - \epsilon_i - w} . \tag{50}$$

We note that this is a polynomial equation of degree K, that is, of degree one lower than the equation for determining G. In particular, for the case of $K = 2$, while we have to solve a cubic equation to determine G, we only have to solve a quadratic equation to determine B!

We will now exploit this fact and study the Hamiltonian $H = H_0 + \varphi$ for $K = 2$; in other words, with no loss of generality we take the symmetric case with the deterministic piece

$$H_0 = \begin{pmatrix} \epsilon & 0 \\ 0 & -\epsilon \end{pmatrix} \tag{51}$$

and take φ to be a Gaussian random perturbation. We can imagine some possible physical applications of this Hamiltonian. For instance, consider electron scattering on impurities in spin-dependent quantum Hall fluids. In the absence of impurity scattering we have a spin-up Landau level separated in energy from the spin-down Landau level, with a Zeeman splitting of 2ϵ. Impurity scattering is represented schematically by φ. A model not precisely of this type, but in the same spirit, was proposed by Hikami, Shirai and Wegner [8] and has been studied by a number of authors [9,10]. The reader can no doubt concoct other possible situations represented, at least schematically, by a Hamiltonian of the type considered here.

For $\epsilon = 0$, the density of states of H is given by the familiar semi-circle law $\rho(\mu) = \frac{1}{2\pi}\sqrt{4-\mu^2}$ which we obtain easily from (3). Note that the half-width of the spectrum is equal to 2. In the other limit, $\epsilon \gg 2$, the density of states decompose into two pieces. At some critical value ϵ_c the two disjoint pieces in the density of states touch and merge into one piece. An interesting question is whether ϵ_c is larger or smaller than 2. A simple physical argument based on level repulsion would suggest that $\epsilon_c < 2$.

Eq. (49) for G becomes

$$2G = \frac{1}{z-\epsilon-G} + \frac{1}{z+\epsilon-G}. \tag{52}$$

While an explicit solution of this cubic equation may be written down, it is rather unwieldy. In contrast, as just explained, B satisfies the quadratic equation

$$zB(z)^2 - (2z^2+1)B(z) + z^3 + (1-\epsilon^2)z = 0, \tag{53}$$

whose solution can of course be immediately written down.

We would now like to determine the end points of the spectrum of H using the algorithm we developed above. We employ the following simple trick. Define the operator $\frac{d}{dz}|$ with the vertical bar indicating that the derivative should not act on B. This operator is clearly useful in light of (48). We now act on (53) with the operator $(z\frac{d}{dz}| - 1)$. The net effect of this operator is to simply replace z^n in (53) by $(n-1)z^n$. We thus obtain a linear equation for B:

$$(2z^2-1)B - 2z^3 = 0. \tag{54}$$

We solve this equation and (53) (or somewhat more simply, the equation obtained by acting with $\frac{d}{dz}|$ on (53)) simultaneously for z and B. The values of B thus obtained are in fact the values of a that we are trying to determine.

Proceeding in this way, we find easily that the four values of a are given by

$$a_+ = \frac{1}{2\sqrt{2}\epsilon}\frac{(4\epsilon^2 - 1 + \sqrt{8\epsilon^2+1})^{3/2}}{\sqrt{8\epsilon^2+1}-1}, \tag{55}$$

$$a_- = \frac{1}{2\sqrt{2}\epsilon}\frac{(4\epsilon^2 - 1 - \sqrt{8\epsilon^2+1})^{3/2}}{\sqrt{8\epsilon^2+1}+1} \tag{56}$$

and $-a_+$ and $-a_-$.

As a first check, we take the limit $\epsilon \to 0$. We find that, as expected, two of the a's, namely $\pm a_-$, become complex and hence unphysical. Indeed, $a_+ \to 2$, and we recover Wigner's classic result for a Gaussian random Hamiltonian.

In the opposite limit with ϵ large, we find

$$a_+ = \epsilon\left(1 + \frac{\sqrt{2}}{\epsilon} + \ldots\right) \tag{57}$$

and

$$a_- = \epsilon\left(1 - \frac{1}{\sqrt{2}\epsilon} + \ldots\right). \tag{58}$$

Thus, in the density of states the width of each of the two disjoint pieces, or bands, is equal to $\sqrt{2} + \frac{1}{\sqrt{2}} = \frac{3}{\sqrt{2}}$. Thus, even when the two bands are far apart, the width is reduced compared to the Gaussian bandwidth by

$$\frac{\text{bandwidth}}{\text{Gaussian bandwidth}} = \frac{3}{4\sqrt{2}}. \tag{59}$$

Finally, the phase transition when the two disjoint bands merge into one occurs when a_- vanishes and becomes complex. This occurs at

$$\epsilon_c = 1. \tag{60}$$

As expected, ϵ_c is less than the unperturbed Gaussian half-width of 2. Interestingly, level repulsion reduces the naive expectation by precisely a factor of two.

Note that for the purpose of this discussion we never had to solve for B explicitly, but the solution of (53) is of course easy enough to write down for the records:

$$B(z) = \frac{2z^2 + 1 + \sqrt{1 + 4\epsilon^2 z^2}}{2z}. \tag{61}$$

The function $B(z)$ has two branch cuts starting at $z_c = \pm i/2\epsilon$. We check easily that the addition law in (19), giving $B(z) = B_0(z) + z$, with B_0 the B function associated with H_0 of course, produces the same result.

Now that we have discussed a specific example, let us now go back to the general case and write (50) as

$$z = \int d\mu \frac{\sigma(\mu)}{B - \mu - z}, \tag{62}$$

where $\sigma(\mu) = \frac{1}{N}\sum_i \delta(\mu - \epsilon_i)$ is the density of states of the deterministic Hamiltonian H_0. Again, with no loss of generality, we have set $m^2 = 1$ in the Gaussian distribution governing φ. The end point a of the spectrum of H is given by the value of B that simultaneously solves (62) and the equation

$$1 = \int d\mu \frac{\sigma(\mu)}{(B - \mu - z)^2}. \tag{63}$$

Note that, as explained before, (63) is obtained by acting with the operator $\frac{d}{dz}|$ on (62).

Given any $\sigma(\mu)$ we can thus determine in principle the end points of the energy spectrum of H. We can for example consider the limit of large randomness, that is, when the spectrum of H_0 is small compared to the scale of the randomness. Expanding the denominators in (62) and (63) and defining $\langle \epsilon^{2k} \rangle \equiv \int d\mu \sigma(\mu) \mu^{2k}$ (assuming for simplicity that $\sigma(\mu)$ is even) we obtain

$$z = \frac{1}{(B - z)} + \frac{\langle \epsilon^2 \rangle}{(B - z)^3} + \frac{\langle \epsilon^4 \rangle}{(B - z)^5} + \ldots \tag{64}$$

and

$$1 = \frac{1}{(B - z)} + \frac{3\langle \epsilon^2 \rangle}{(B - z)^4} + \frac{5\langle \epsilon^4 \rangle}{(B - z)^6} + \ldots \tag{65}$$

In the limit $H_0 \to 0$ or large randomness, we have $z = \frac{1}{B-z} = \pm 1$ (and thus the end points $a = \pm 2$, as expected.) It is easy to solve these equations to any desired power. We obtain to $O(H_0^4)$ that the width of the spectrum is given by

$$a = 2 + \langle \epsilon^2 \rangle - \tfrac{9}{4}\langle \epsilon^2 \rangle^2 + \langle \epsilon^4 \rangle + \ldots \tag{66}$$

A simple check shows that indeed, if we expand a_+ in (55), we find

$$a_+ = 2 + \epsilon^2 - \tfrac{5}{4}\epsilon^4 + \ldots \tag{67}$$

This is consistent with (66) since in this simple model $\langle \epsilon^2 \rangle^2 = \langle \epsilon^4 \rangle = \epsilon^4$.

5. Analytic structure

For the problem studied here B is simpler than G, but in other problems B is unfortunately more complicated. For instance, for a much studied class of probability distributions, the so-called trace class, defined by (2), and with V taken to be an even polynomial for simplicity, $G(z)$ is given by [5], as mentioned earlier,

$$G(z) = \frac{1}{2}\left[V'(z) - P(z)\sqrt{z^2 - a^2}\right], \tag{68}$$

where the polynomial P and the end point a are determined completely by the condition $G(z) \to 1/z$ as $z \to \infty$. Applying our rule of substituting $G \to w$ and $z \to B$, we obtain the equation

$$2w = V'(B) - P(B)\sqrt{B^2 - a^2} \tag{69}$$

which determines B.

For V of degree $2s$ it is easy to see that B satisfies a polynomial equation of degree $2s-1$. Indeed, moving $V'(B)$ in (69) to the left-hand side and squaring, we can rewrite (69) as

$$wV'(B) - w^2 - Q(B) = 0, \tag{70}$$

where $4Q(B) \equiv V'(B)^2 - P(B)^2(B^2 - a^2)$ is a polynomial of degree $2s-1$ in B. (This is most easily seen by noting that $G(B) \to 1/B$ for large B by definition.)

We would like to study the analytic structure of B but we have not been able to make general and complete statements. One limited statement applies to the trace class just discussed. It is easy to show then that the branch cuts of $B(w)$ is of the square root type. Let $B(w)$ has a cut starting at $w*$ and write $B* = B(w*)$. At the tip of the cut,

$$\left.\frac{dB(w)}{dw}\right|_{w*} = \infty \tag{71}$$

or

$$\left.\frac{dG(B)}{dB}\right|_{B*} = 0. \tag{72}$$

Note that these equations are the duals of the ones in (46) and (48). Expanding, we find immediately that

$$B = B* + \left(\left. \frac{d^2Q}{dB^2} \right|_B * \right)^{-1/2} 4(w - w*)^{1/2} + \dots \tag{73}$$

For the case studied in the previous section with H_0 given as in (51) we see from the explicit form given in (61) that $B(w)$ indeed has a square root cut starting at $w* = \pm i/2\epsilon$, in agreement with the general analysis given here.

6. Non-abelian central limit theorem

Gauss proved that if we add K random numbers x_i, $i = 1, 2, \dots, K$, with x_i taken from the probability distribution $P_i(x_i)$, then the normalized sum $s = \frac{1}{\sqrt{K}} \sum_i x_i$ follows the Gaussian distribution in the limit K tending to infinity. This result plays an important role in physics and mathematics, and accounts for the ubiquitous appearance of the Gaussian distribution.

What if the variables x_i do not commute? In particular, suppose that instead of real numbers x_i, we have N by N random matrices φ_i, $i = 1, 2, \dots, K$, taken from the probability distributions

$$P_i(\varphi_i) = \frac{1}{Z_i} e^{-N \operatorname{tr} V_i(\varphi_i)}. \tag{74}$$

Does the normalized sum of all these matrices $\varphi_T \equiv \frac{1}{\sqrt{K}} \sum_i \varphi_i$ follow the Gaussian distribution in the large K limit? Intuitively, it seems that this ought to be the case. (Note that there exists another commonly considered class of random matrices, in which the element of the random matrices is each taken from a probability distribution (the same for each element). For this class, which we referred to as the Wigner class in our earlier work [12], the proposed theorem follows immediately from the usual abelian central limit theorem. Here we are speaking of the trace classes defined by (74).)

As it turns out, it is not difficult to generalize one of the standard proofs of Gauss' theorem to matrices. We will give this proof below. However, it would seem that the algorithm developed here for adding random matrices is almost tailor made to address this question of whether $\varphi_T \equiv \frac{1}{\sqrt{K}} \sum_i \varphi_i$ follows the Gaussian distribution. It is mildly amusing to see how the Wigner semi-circle law emerges naturally.

As we mentioned earlier, when we add two random matrices, in general it is difficult to determine explicitly the resulting $G(z)$ for the sum of the two matrices. What we hope for here is that the large K limit will bring considerable simplification. This is indeed the case. To keep the formulas simple, let us again take V_i to be even. We expect that our conclusions can be easily generalized. In this case, $G_i(z)$ is an odd function and hence $B_i(z)$ is also an odd function. Thus, we write $B_i(z) = 1/z + z b_i(z)$ with $b_i(z)$ an even function. From the law of addition (19) given in this paper, we learn that the function $B(z)$ associated with the unknown $G(z)$ is given by

$$B(z) = \frac{1}{z} + zb(z) , \tag{75}$$

where

$$b(z) = \sum_{i=1}^{K} b_i(z) . \tag{76}$$

Thus, we obtain the simple result that the Green's function $G(z)$ associated with $\sum_i \varphi_i$ is determined by solving

$$\frac{1}{G(z)} + b(G(z))G(z) = z . \tag{77}$$

We are however interested in the normalized sum $\varphi_T \equiv \frac{1}{\sqrt{K}} \sum_i \varphi_i$. Define the corresponding Green's function as

$$G_T(z) \equiv \langle \frac{1}{N} \mathrm{tr} \frac{1}{z - \frac{1}{\sqrt{K}} \sum_i \varphi_i} \rangle = \sqrt{K} G(\sqrt{K} z) .$$

Letting $z \to \sqrt{K} z$ in (77), we find that $G_T(z)$ is determined by

$$\frac{1}{G_T(z)} + \left[\frac{1}{K} \sum_i^K b_i \left(\frac{G_T(z)}{\sqrt{K}} \right) \right] G_T(z) = z . \tag{78}$$

Note that the argument of b_i in (78) is $G_T(z)/\sqrt{K}$. For finite K, (78) is hopelessly complicated. However, as $K \to \infty$ we see that it simplifies rather naturally to

$$\frac{1}{G_T(z)} + \sigma^2 G_T(z) = z , \tag{79}$$

with

$$\sigma^2 \equiv \frac{1}{K} \sum_i^K b_i(0) . \tag{80}$$

Solving this quadratic equation, we find immediately that

$$G_T(z) = \frac{1}{2} \left(z - \sqrt{z^2 - 4\sigma^2} \right) \tag{81}$$

and thus Wigner's semi-circle law for the density $\rho(\mu) = \frac{1}{2\pi\sigma^2} \sqrt{4\sigma^2 - \mu^2}$.

What is $b_i(0)$? We note that with $G_i(z) \to \frac{1}{z} + \frac{\sigma_i^2}{z^3} + \ldots$ for large z it is easy to show that $b_i(0) = \sigma_i^2$. Thus, not only do we obtain the Wigner semi-circle law, we learn that $\sigma^2 = \frac{1}{K} \sum_i^K \sigma_i^2$. We should remark that in general σ_i^2 is not directly related to the width of the spectrum of eigenvalues of φ_i, as the reader can easily check by using the explicit formulas given in footnote 1.

While we have proved that the density of eigenvalues of φ_T satisfies the semi-circle law, we cannot yet conclude that the probability distribution of φ is Gaussian. The reason is that for matrices in the Wigner class, even when the distribution is not Gaussian, the

corresponding density of eigenvalues still satisfies the semi-circle law, as is well known. (For a simple proof based on a renormalization group inspired approach, see Ref. [12].) For the trace class, on the other hand, it is true that if the density satisfies the semi-circle law, then the distribution of the matrices is indeed Gaussian [5]. However, we see no reason that φ_T would belong to the trace class, and not to a more involved class of probability distributions such as those discussed in [12].

As mentioned above, it is not difficult to extend one of the usual proofs of the central limit theorem to the case of matrices. The distribution for the normalized sum matrix $\varphi \equiv \frac{1}{\sqrt{K}} \sum_i \varphi_i$ is given by (here we omit the subscript T)

$$P(\varphi) = \left(\prod_i^K \int d\varphi_i P(\varphi_i) \right) \delta \left(\varphi - \frac{1}{\sqrt{K}} \sum_i \varphi_i \right)$$
$$= \int dt \left(\prod_i^K \int d\varphi_i P(\varphi_i) \right) \exp \left(\frac{i}{\sqrt{K}} \sum_i^K \operatorname{tr} t\varphi_i \right) \exp(-i \operatorname{tr} t\varphi) \ . \tag{82}$$

The integral over φ_i can be done in the large K limit:

$$\int d\varphi_i \, P(\varphi_i) \exp \left(\frac{i}{\sqrt{K}} \sum_i \operatorname{tr} t\varphi_i \right) = 1 - \frac{1}{2K} \int d\varphi_i \, P(\varphi_i) \operatorname{tr} t\varphi_i \operatorname{tr} t\varphi_i + O\left(\frac{1}{K^2} \right)$$
$$= 1 - \frac{\sigma_i^2}{2KN} \operatorname{tr} t^2 + O\left(\frac{1}{K^2} \right) , \tag{83}$$

where we have defined $\sigma_i^2 = \int d\varphi_i P(\varphi_i) \frac{1}{N} \operatorname{tr} \varphi_i^2$. Reexponentiating and integrating over t we obtain the desired result that $P(\varphi)$ is proportional to $\exp[-(N/4\sigma^2) \operatorname{tr} \varphi^2]$.

Acknowledgements

I thank the organizers of the workshop on "Operator Algebra Free Products and Random Matrices" held at the Fields Institute, March 1995, for inviting me to talk about my work. I had the opportunity there to learn from mathematicians, in particular U. Haagerup and D. Voiculescu, about recent developments. I am especially grateful to M. Douglas for many instructive discussions and for telling me how to translate from the language used by mathematicians to the language used by physicists. Finally, I would like to thank E. Brézin, my collaborator over the last few years, for countless helpful discussions about this and other topics in random matrix theory. The results of Sections 4–6 were worked out together with him. This work was supported in part by the National Science Foundation under Grant No. PHY89-04035.

Note added

For an application of the formalism discussed in this paper to particle physics, see the recent preprint of Nowak et al. [23].

References

[1] E. Wigner, Can. Math. Congr. Proc. (University of Toronto Press) p. 174, and other papers reprinted in Ref. [2].
[2] C.E. Porter, Statistical Theories of Spectra: Fluctuations (Academic Press, New York, 1965).
[3] M.L. Mehta, Random Matrices (Academic Press, New York, 1991).
[4] See for instance, Two Dimensional Quantum Gravity and Random Surfaces, ed. D.J. Gross and T. Piran (World Scientific, Singapore, 1992).
[5] E. Brézin, C. Itzykson, G. Parisi and J.B. Zuber, Commun. Math. Phys. 59 (1978) 35.
[6] E. Brézin and A. Zee, Nucl. Phys. B 402 [FS] (1993) 613.
[7] J. Ambjørn, Quantization of geometry, *in* Les Houches (1990), ed. J. Dalibard et al. (North-Holland, Amsterdam);
J. Ambjørn, J. Jurkiewicz and Yu.M. Makeenko, Phys. Lett. B 251 (1990) 517.
[8] C.W.J. Beenakker, Nucl. Phys. B 422 [FS] (1994) 515.
[9] J. D'Anna, E. Brézin and A. Zee, Nucl. Phys. B 443 [FS] (1995) 433.
[10] P.J. Forrester, Nucl. Phys. B 435 [FS] (1995) 421.
[11] T.S. Kobayakawa, Y. Hatsugai, M. Kohmoto, and A. Zee, Phys. Rev. E 51 (1995) 5365.
[12] E. Brézin and A. Zee, C. R. Acad. Sci. (Paris) t.317 II (1993) 735.
[13] E. Brézin and A. Zee, Phys. Rev. E 49 (1994) 2588.
[14] E. Brézin, S. Hikami and A. Zee, Phys. Rev. E 51 (1995) 5442.
[15] D.V. Voiculescu, K.J. Dykema and A. Nica, Free Random Variables (AMS, Providence, RI, 1992).
[16] U. Haagerup, private communication and to be published.
[17] Lectures at the workshop on Operator Algebra Free Products and Random Matrices, Fields Institute (March 1995), to appear in the Proceedings.
[18] M. Douglas, Rutgers preprint, hep- th/9409098 (1994).
[19] R. Gopakumar and D.J. Gross, Princeton preprint PUPT-1520 (1994).
[20] E. Brézin and A. Zee, Nucl. Phys. B 441 [FS] (1995) 409.
[21] E. Brézin and A. Zee, Nucl. Phys. B 424 [FS] (1994) 435.
[22] M. Douglas and M. Li, Rutgers preprint (1995).
[23] M.A. Nowak, G. Papp and I. Zahed, hep-ph/9603348.
[24] B. Eynard, C. Kristjansen, Nucl. Phys. B 466 [FS] (1996) 463.
[25] G. 't Hooft, Nucl. Phys. B 72 (1974) 461.
[26] L.A. Pastur, Theor. Math. Phys. 10 (1972) 67.