Journal of Statistical Physics, Vol. 98, Nos. 1/2, 2000

The Dilute Potts Model on Random Surfaces

P. Zinn-Justin[1]

Received April 23, 1999

We present a new solution of the asymmetric two-matrix model in the large-N limit which only involves a saddle point analysis. The model can be interpreted as Ising in the presence of a magnetic field, on random dynamical lattices with the topology of the sphere (resp. the disk) for closed (resp. open) surfaces; we elaborate on the resulting phase diagram. The method can be equally well applied to a more general $(Q+1)$-matrix model which represents the dilute Potts model on random dynamical lattices. We discuss in particular duality of boundary conditions for open random surfaces.

KEY WORDS: Random surfaces; random dynamical lattices; matrix models; Potts model.

1. INTRODUCTION

The study of various multi-matrix models in the large N limit is motivated by their interpretation as statistical lattice models on random surfaces. The one-matrix model(1) already describes the summation over random surfaces, but to put "matter" on the surface, several matrices are required.[2] The simplest such model is the two-matrix model, which has the following partition function(3, 5)

$$Z(\alpha_0, \beta_0, \gamma) = \iint \mathrm{d}A \, \mathrm{d}B \, \mathrm{e}^{N \operatorname{tr}[-1/2(A^2+B^2) + (\alpha_0/3) A^3 + (\beta_0/3) B^3 + (1/\gamma) AB]} \tag{1.1}$$

where A and B are $N \times N$ hermitean matrices. In the large N limit, this generates triangulated surfaces with the spherical topology, on which spins live (the two matrices A and B correspond to spins up and down),(7) thus

[1] Department of Physics and Astronomy, Rutgers University, Piscataway, New Jersey 08854-8019; e-mail: pzinn@physics.rutgers.edu.

[2] Except the Kazakov multicritical points(2) which correspond to non-unitary matter.

0022-4715/00/0100-0245$18.00/0

reproducing the Ising model on random surfaces. Schematically, α_0 and β_0 play the roles of both magnetic field H and cosmological constant ($\alpha_0/\beta_0 = \exp(2H/T)$), while γ is related to the temperature T by $\gamma = \exp(1/T)$.

This model has been solved using orthogonal polynomials,[5] but, strangely, the simplest tool available, the saddle point method, which works so well for the one-matrix model, has not been used. It is commonly assumed that this method does not work in the two-matrix model, the usual argument being the following: after diagonalization of A and B and use of the Itzykson–Zuber–Harish Chandra formula,[4, 3] we are left with an integral over the eigenvalues:

$$Z(\alpha_0, \beta_0, \gamma) = \iint \prod_i da_i\, db_i\, \Delta[a_i]\, \Delta[b_i]\, \det_{i,j}[e^{N(1/\gamma)\, a_i b_j}] \times e^{N \sum_i [-1/2(a_i^2 + b_i^2) + (\alpha_0/3)\, a_i^3 + (\beta_0/3)\, b_i^3]} \tag{1.2}$$

where $\Delta[\cdot]$ is the Van der Monde determinant. Then, one uses symmetry of permutation of the eigenvalues to reduce this expression to:

$$Z(\alpha_0, \beta_0, \gamma) = \iint \prod_i da_i\, db_i\, \Delta[a_i]\, \Delta[b_i] \times e^{N \sum_i [-1/2(a_i^2 + b_i^2) + (\alpha_0/3)\, a_i^3 + (\beta_0/3)\, b_i^3 + (1/\gamma)\, a_i b_i]} \tag{1.3}$$

At this stage, a large N saddle point analysis shows that there is a continuous infinity of saddle points and it is very difficult to derive anything from it. Of course, the problem stems from the transformation of (1.2) to (1.3): since we have broken the symmetry of permutation of the eigenvalues, each ordering $a_{\sigma(1)} < \cdots < a_{\sigma(N)}$ of the eigenvalues leads to a different saddle point, so that we have $N!$ saddle points, which causes trouble as $N \to \infty$. The problem does not exist at the level of (1.2) and in fact, as we shall show, there is a well-defined saddle point to it; we shall then explain how to determine it, which will result in a very compact and elegant way of expressing the resolvent(s) of the model. A similar result could be obtained for the two-matrix model with quartic vertices, but we shall not choose to do so here.

We shall then generalize our results to the dilute Q-states Potts model on random surfaces, which is defined by (see ref. 22 for a somewhat similar definition on a flat lattice)

$$Z_Q(\alpha_0, \beta_0, \gamma) = \iint dA \prod_{q=1}^{Q} dB_q \times e^{N\, \mathrm{tr}[-1/2(A^2 + \sum_{q=1}^{Q} B_q^2) + (\alpha_0/3)\, A^3 + \sum_{q=1}^{Q} ((\beta_0/3)\, B_q^3 + (1/\gamma)\, A B_q)]} \tag{1.4}$$

This model describes the following statistical model on random surfaces: each vertex of the surface is either unoccupied (represented by matrix A) or occupied by a spin in Q possible states (matrices B_q). The matrices B_q only interact via the matrix A (in particular configurations with adjacent vertices in distinct spin states appear only at order $1/\gamma^2$). Again, α_0 and β_0 are cosmological constants and control the dilution (i.e., density of unoccupied sites), and γ is related to the inverse temperature. The two-matrix model (1.1) is the particular case $Q=1$.

From now on we shall redefine the couplings of the model and rescale the fields so that the partition function can be rewritten

$$Z_Q(\alpha, \beta, \gamma) = \int dA\, e^{N \operatorname{tr}[-(\gamma/2) A^2 + (\alpha/3) A^3]} \left(\int dB\, e^{N \operatorname{tr}[-(\gamma/2) B^2 + (\beta/3) B^3 + AB]} \right)^Q \tag{1.5}$$

where $\alpha = \alpha_0 \gamma^{3/2}$, $\beta = \beta_0 \gamma^{3/2}$. The main physical quantities of the model are the resolvents, which are defined by

$$\begin{aligned} \omega_A(a) &= \left\langle \frac{1}{N} \operatorname{tr} \frac{1}{a-A} \right\rangle \\ \omega_B(b) &= \left\langle \frac{1}{N} \operatorname{tr} \frac{1}{b-B} \right\rangle \end{aligned} \tag{1.6}$$

where the large N limit is implied and a, b are complex numbers. $\omega_A(a)$ and $\omega_B(b)$ are generating functions of averages of the form $\langle \operatorname{tr} A^n \rangle$ and $\langle \operatorname{tr} B^n \rangle$, but they are also important from the diagrammatic point of view: they correspond to sums over connected surfaces with a boundary ("loop functions"), the parameters a or b playing the role of boundary cosmological constant. In the large N limit, these surfaces have the topology of a disk. The difference between $\omega_A(a)$ and $\omega_B(b)$ lies in the *boundary conditions*: for $\omega_A(a)$ (resp. $\omega_B(b)$), there are only matrices A (resp. B) on the boundary. The large n asymptotics of $\operatorname{tr} A^n$ and $\operatorname{tr} B^n$ (i.e., surfaces with large boundary) are dominated by the singularities of the corresponding resolvent. These singularities are also relevant for the physics in the bulk: if g is the exponent of this singularity, then the central charge c of the critical model is given by[18, 19]

$$c = 1 - 6(\sqrt{g} - 1/\sqrt{g})^2 \tag{1.7}$$

These resolvents will therefore play a central role in our analysis, and our goal will be to find exact expressions for them. Let us remark that when the dilution is turned off ($\alpha = 0$), one can perform the gaussian integration over A and we are brought back to the usual Q-states Potts model on random surfaces.[15] $\omega_B(b)$ is then the standard resolvent: it corresponds to boundary conditions of the Potts model where all the spins on the boundary are in a given state. $\omega_A(a)$ is not a natural resolvent, because the sites on the boundary are unoccupied, whereas they cannot be so in the bulk. However, once dilution it turned on, the two resolvents $\omega_A(a)$ and $\omega_B(b)$ should be treated *on equal footing*; they will correspond, as we shall explain later, to boundary conditions which are "dual" to each other.

The plan of the article therefore goes as follows: first we shall analyze in Section 2 the integral over B which is common to the models (1.5) for arbitrary Q, then discuss the cases $Q = 1, 2$ and $3, 4$ (though in (1.5) the parameter Q can take arbitrary real values, for simplicity we only consider here integer values) in Sections 3, 4, 5 and finally conclude in Section 6. The appendix presents a comparison with the method of orthogonal polynomials (in the case of the two-matrix model).

2. THE EXTERNAL FIELD PROBLEM

Let us first consider part of the model only: if the matrix A is held fixed in (1.5), we are left with the problem of one matrix in an external field[13, 14]:

$$\Xi(A) = \int \mathrm{d}B \, \mathrm{e}^{N \operatorname{tr}[-V(B) + AB]} \tag{2.1}$$

where V is a polynomial potential. We shall show how the saddle point equations expressed in ref. 11 allow indeed to calculate this matrix integral, and how in the case of the cubic potential $V(b) = -(\beta/3)\, b^3 + (\gamma/2)\, b^2$, they in fact reproduce the solution[14] without the use of any partial differential equations.

By $U(N)$-invariance $\Xi(A)$ only depends on the eigenvalues a_i of A:

$$\Xi[a_i] = \int \prod_i \mathrm{d}b_i \, \Delta[b_i] \, \frac{\det_{i,j} [\mathrm{e}^{N a_i b_j}]}{\Delta[a_i]} \, \mathrm{e}^{-N \sum_i V(b_i)} \tag{2.2}$$

We shall now show how to calculate in the large N limit the logarithmic derivatives of Ξ with respect to a_i (Ξ then follows by simple integration).

We assume that the density of eigenvalues of A becomes smooth in the large N limit, and denote it $\rho_A(a)$; for the sake of simplicity only, its support will be taken to be of the form of a single interval $[a_1, a_2]$. It is related to the resolvent of A by

$$\omega_A(a) = \int_{a_1}^{a_2} \frac{\mathrm{d}a' \rho_A(a')}{a - a'} \tag{2.3}$$

which is an analytic function everywhere except on the support of A, where it has a cut (leading, if $\rho_A(a)$ is smooth, to other sheets) which we call the physical cut.

We can now write[11]

$$\frac{1}{N} \frac{\partial}{\partial a_i} \log \Xi[a_i] = b(a_i) - \omega_A(a_i) \tag{2.4}$$

where $b(a)$ is an analytic (multi-valued) function which has the same physical cut as $\omega_A(a)$. The particular sheet corresponding to the value of $b(a)$ in (2.4) is called the physical sheet. We see that $b(a)$ is the quantity we need to compute.

In order to do so, we must write down the saddle point equation of (2.1), which we shall do carefully. First we introduce symmetrically the resolvent $\omega_B(b)$ of B, with a physical cut $[b_1, b_2]$ (i.e., the eigenvalues of B fill the interval between b_1 and b_2):

$$\omega_B(b) = \int_{b_1}^{b_2} \frac{\mathrm{d}b' \, \rho_B(b')}{b - b'}$$

Taking the logarithmic derivative with respect to the b_i results in the appearance of the function $a(b)$, which has the same cut as the resolvent $\omega_B(b)$ of B, and is the *functional inverse* of $b(a)$ (see Appendix 1 of ref. 11). The saddle point equations then read:

$$\tfrac{1}{2}(\omega_B(b + i0) + \omega_B(b - i0)) + \tfrac{1}{2}(a(b + i0) + a(b - i0)) = V'(b) \quad b \in [b_1, b_2]$$

Using the fact that $a(b)$ and $\omega_B(b)$ have the same cut, this equation can be analytically continued:

$$\omega_B(b) + a_\star(b) = V'(b) \tag{2.5}$$

where $a_\star(b)$ is the value of $a(b)$ on the other side of the physical cut of B. Before going further, let us give the physical significance of $a(b)$. Formally, the saddle point equation (2.5) can be written down: $(\mathrm{d}/\mathrm{d}b)(\delta S/\delta \rho_B(b)) = 0$, where S is the action, and the derivative $\mathrm{d}/\mathrm{d}b$ takes care of the normalization

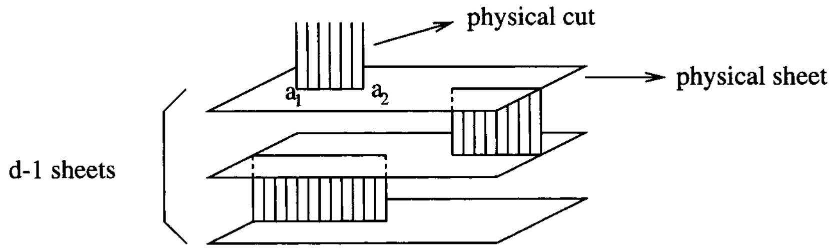

Fig. 1. Analytic structure of $b(a)$ in the external field problem with a polynomial potential of degree d.

condition $\int \mathrm{d}b\, \rho_B(b) = 1$. Now $\delta S/\delta \rho_B(b)$ is nothing but the *effective potential* $V_{\text{eff}}(b)$ for the eigenvalues b_i, which must be the sum of three terms: the potential $V(b)$, the effective interaction among the b_i, and the effective interaction between the b_i and the a_i. Indeed, we find explicitly that

$$V_{\text{eff}}(b) = \begin{cases} \text{const} & b \in [b_1, b_2] \\ V(b) - \int^b [\omega_B(b) + a(b)]\, \mathrm{d}b & b \notin [b_1, b_2] \end{cases}$$

(note that it is $a(b)$ which appears and not $a_\star(b)$ because we are outside the cut). Therefore $a(b)$ can be interpreted as the derivative of the effective potential for the action of the a_i on the b_i.[3]

Let us show now how to handle the equation (2.5): it can be rewritten

$$\omega_B(b(a)) + a = V'(b(a)) \tag{2.6}$$

where it is understood that (2.6) is a relation between multi-valued functions. To use this equation, we now assume that V is a polynomial of degree d, and look at the behavior of $b(a)$ as $a \to \infty$. Under minimal assumptions on the analytic structure of $b(a)$ (see below), on all cuts connected to the physical sheet without crossing the physical cut, $b(a) \to \infty$ as $a \to \infty$. Therefore $b(a)^{d-1} \sim a$ as $a \to \infty$, which leads us to the "minimal" conjecture for the analytic structure of $b(a)$ which is depicted by Fig. 1 (a similar structure was found for large N characters in ref. 17). All the cuts except the physical cut go to infinity, and we assume that they do not cross the physical cut.[4]

[3] Symmetrically, $b(a)$ will be the derivative of the effective potential for the action of the b_i on the a_i.

[4] This corresponds to the strong coupling phase of ref. 13. In the present context, it simply means we remain below the continuum limit surface.

At this point we have complete knowledge of the analytic behavior of $b(a)$ (the physical cut is determined by the fact that it is identical to that of $\omega_A(a)$, which is known by definition: $b(a-i0)-b(a+i0)=2\pi i\rho_A(a)$), and this is enough to compute it.

More explicitly, by a change of variables of the form: $a\equiv P(z)$, where P is a polynomial of degree $d-1$ whose critical values are the branching points, we can remove all the cuts at infinity; then $b(z)$ has a single cut, the physical cut, and it can be expressed as:

$$b(z)=c_1 z+c_2+\int_{z(a_1)}^{z(a_2)} \frac{dz' \, \rho_A(a(z'))}{z-z'} \tag{2.7}$$

where the constants c_1 and c_2 are easily determined by asymptotics at infinity.

The simplest non-trivial case is the cubic case, in which we find only two sheets (the physical sheet and an extra one) if we do not cross the physical cut, and the equation (2.7) reduces to the known solution.[14] Since this is the case that is of interest to us, let us carry out explicitly the procedure outlined above. From Eq. (2.6) with $V(b)=-(\beta/3)\, b^3+(\gamma/2)\, b^2$, we obtain that there are two sheets $b_\pm(a)$ (the physical sheet being by definition $b_+(a)$) on which as $a\to\infty$,

$$b_\pm(a)=\pm(-a/\beta)^{1/2}+\frac{\gamma}{2\beta}\pm\frac{\gamma^2}{8\beta^{3/2}}(-a)^{-1/2}+\frac{1}{2a}+O(a^{-3/2}) \tag{2.8}$$

where we have used $\omega_B(b)=1/b+O(1/b^2)$.[5] After the change of variables $z^2=a_3-a$ (where a_3 is the branch point of the semi-infinite cut), we find that

$$b(z)=\frac{z}{\beta^{1/2}}+\frac{\gamma}{2\beta}+\frac{\gamma^2-4a_3\beta}{8\beta^{3/2}}\frac{1}{z}-\frac{1}{2z^2}+O(1/z^3)$$

In particular the constants in (2.7) are given by $c_1=\beta^{-1/2}$ and $c_2=\gamma/2\beta$.

Note however that in the next paragraphs, we shall not need to make the explicit change of variables $a\to z$. It is known that this maps the saddle point equations of the Q-states Potts model onto those of the $O(n)$ model;[16] even though this correspondence allows in principle to solve the Q-states Potts model using the general solution of the $O(n)$ model, it does not mean that the models are physically equivalent (the phase diagrams are

[5] Had we used an expansion $\omega_B(b)=\sum_{n=0}^{\infty} B_n/b^{n+1}$ where $B_n\equiv\langle(1/N)\,\mathrm{tr}\, B^n\rangle$, we would have obtained and expansion at arbitrary order of $b_\pm(a)$ in terms of the B_n.

different, the critical models do not have the same central charge,[6] etc.) and it is not convenient for our purposes. Instead, we shall directly consider $b(a)$ in its normal parametrization and use the information on its analytic structure discussed above.

3. THE $Q=1$ DILUTE POTTS MODEL (ISING WITH MAGNETIC FIELD)

Let us now apply the results of the previous paragraph to the two-matrix model. The philosophy will be the same in this case as in the more general Q-states Potts model with $0<Q<4$: we have two saddle point equations, one coming from A, the other one from B. These equations involve functions $a(b)$ and $b(a)$ which satisfy a functional inversion relation; therefore, by inverting one of them, we obtain two relations for $b(a)$. Recombining them, we deduce a polynomial equation satisfied by $b(a)$.

The saddle point equation of the eigenvalues of B has already been analyzed in the previous section; as we have shown, it gives the behavior of $b(a)$ at infinity and fixes its analytic structure (Eq. (2.8)). We have found two sheets, the physical sheet $b_+(a)$, and $b_-(a)$ which is connected to it by a semi-infinite cut.

Remembering that the full partition function takes the form

$$Z_1(\alpha, \beta, \gamma) = \iint \prod_i \mathrm{d}a_i\, \mathrm{d}b_i\, \Delta[a_i]\, \Delta[b_i]\, \det_{i,j}[\mathrm{e}^{N a_i b_j}] \times \mathrm{e}^{N \sum_i [-(\gamma/2)(a_i^2+b_i^2) + (\alpha/3)\, a_i^3 + (\beta/3)\, b_i^3]}$$

in terms of the eigenvalues of A and B, we now write the analytically continued saddle point equation for the eigenvalues of A:

$$\omega_A(a) + b_\star(a) - \gamma a + \alpha a^2 = 0 \tag{3.1}$$

$b_\star(a)$ is a third sheet of $b(a)$, connected to the physical sheet $b_+(a)$ by the physical cut, and which according to (3.1), has no other cut than the physical cut. This means that $b(a)$ has exactly three sheets, and we therefore make the Ansatz that $b(a)$ is the solution of a third degree equation in a. The coefficients of this polynomial are symmetric functions of the different sheets; for example, using the elementary identity $b_+(a) + b_-(a) = \omega_A(a) + (\gamma/\beta)$ (coming from the analytic structure of $b(a)$) and (3.1), we find that $b_+(a) + b_-(a) + b_\star(a) = -\alpha a^2 + \gamma a + (\gamma/\beta)$. For other symmetric functions we must use the finite expansion (2.8), which implies that not all of the

[6] Technically, this comes from the fact that the mapping $a \mapsto z$ changes by a factor of 2 the critical exponent g related to the central charge by Eq. (1.7).

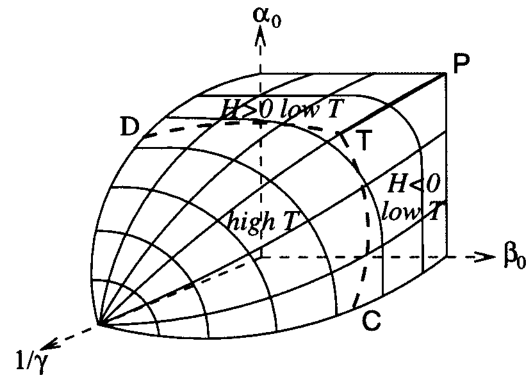

Fig. 2. Schematic phase diagram of the Ising model with magnetic field on random surfaces. PT is the zero magnetic field low temperature phase transition line; T is the critical point of the model; CT resp. DT) is a boundary phase transition line for $b(a)$ (reps. $a(b)$).

coefficients of the polynomial can be found explicitly. We find that there are three remaining unknowns called x, y, z (which can be reexpressed in terms of the three first moment of b, cf. footnote 4). The final equation can be cast in the nicely symmetric form:

$$\alpha\beta a^2 b^2 + \alpha a^3 - \gamma\alpha a^2 b - \gamma\beta b^2 a + \beta b^3 - \gamma a^2 + (\gamma^2 + 1 - \alpha\beta)\, ab - \gamma b^2 + xa + yb + z = 0 \tag{3.2}$$

The constants x, y, z are determined by imposing that $a(b)$ and $b(a)$ have the appropriate analytic structure: when solving (3.2) for b, the discriminant is a 9th degree polynomial in a, and we must impose that 3 zeroes are double zeroes in order to have only three branch points.[7] This gives three algebraic equations for x, y, and z which are therefore functions of α, β, γ.

One can show that, in the absence of magnetic field ($\alpha = \beta$), the equation (3.2) is equivalent to the equation for the resolvent found in refs. 8, 9, and 10 using more complicated methods; however, (3.2) displays explicitly the $\mathbb{Z}_2$ symmetry $a \leftrightarrow b$, $\alpha \leftrightarrow \beta$, whereas it is not obvious in refs. 8, 9, and 10. This explicit symmetry is due to the fact that we are not trying to write down an equation for $\omega_A(a)$ directly (or $\omega_B(b)$), but rather for $b(a)$, which differs by a polynomial part $-\gamma a + \alpha a^2$.

Figure 2 shows the resulting topology for the phase diagram of the model. We shall now explain it by briefly analyzing Eq. (3.2) in two particular cases.

[7] Note that this constraint is much stronger than the one given in ref. 10 ("one cut hypothesis"), which in fact would not be enough to fix the three unknowns.

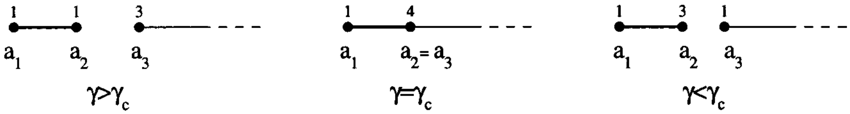

Fig. 3. Cut structure of $b(a)$ for $\alpha = 0$ (infinite magnetic field) on the continuum limit surface. The numbers represent the multiplicity of the zeroes of the discriminant. Note that the two cuts never intersect each other (the choice of cuts is dictated by analytic continuation from the gaussian model).

If we first remove the dilution ($\alpha = 0$, which in the Ising language, corresponds to an infinite magnetic field), this model is simply the one-matrix model, and we expect no physics at all; in fact, integrating over A shows that there is only one parameter in the model, $\beta/(\gamma - 1/\gamma)^{3/2}$. The continuum limit, which is attained for

$$\frac{\beta^2}{(\gamma - 1/\gamma)^3} = \frac{1}{12\sqrt{3}}$$

corresponds to a $c = 0$ theory, and there seems to be no critical point. However, this is *wrong*, because even though the bulk theory is pure gravity with a coupling $\beta/(\gamma - 1/\gamma)^{3/2}$, $b(a)$ represents a non-trivial loop function, and the corresponding boundary operator depends explicitly on γ. Indeed we find that the standard resolvent $\omega_B(b)$, or equivalently $a(b)$ (cf. Eq. (2.5)) always has the singularity

$$a - a_{\text{reg}} \sim (b - b_\star)^{3/2}$$

of the usual pure gravity loop function, but that $b(a)$ (or $\omega_A(a)$) undergoes a phase transition! This can be seen in the behavior of the branch points of $b(a)$ as one varies γ (Fig. 3). There is a critical point[8] (point C of Fig. 2)

$$\gamma_C^2 = 1 + 2\sqrt{3} \qquad \beta_C^2 = 2\gamma_C^{-3}$$

where the two cuts merge, and the result is that

$$b - b_{\text{reg}} \sim \begin{cases} (a - a_2)^{3/2} & \gamma < \gamma_C \\ (a - a_2)^{2/3} & \gamma = \gamma_C \\ (a - a_2)^{1/2} & \gamma > \gamma_C \end{cases} \tag{3.3}$$

For $\gamma > \gamma_C$ there is again a $(a - a_3)^{3/2}$ singularity, but a_3 does not belong to the physical cut.

[8] This is the same critical point found in a related percolation model in ref. 23, where a much more involved analysis shows explicitly the collision of the cuts (found here from the analysis of the discriminant of a third degree equation).

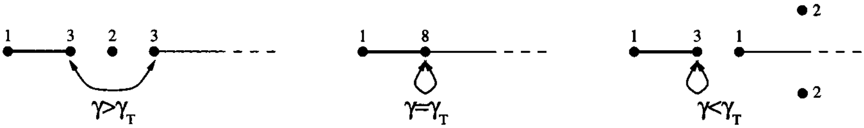

Fig. 4. Cut structure of $b(a)$ for $\alpha = \beta$ (zero magnetic field) on the continuum limit surface. The numbers represent the multiplicity of the zeroes of the discriminant. The exchange of A and B, represented by the arrows, acts non-trivially on the branch points in the low temperature phase ($\gamma > \gamma_T$), but not in the high temperature phase ($\gamma < \gamma_T$).

Let us insist that the corresponding theory on the sphere (i.e., the free energy of the matrix model) only depends on the combination $\beta/(\gamma - 1/\gamma)^{3/2}$ and is always pure gravity; but there is nonetheless a phase transition, at the boundary of the surface, which is signalled by a change of analytic behavior of $b(a)$ from $\gamma < \gamma_C$ (high temperature) to $\gamma > \gamma_C$ (low temperature) —or equivalently, a change in the asymptotics of $\langle \mathrm{tr}\, A^n \rangle$ as $n \to \infty$. The situation might seem trivial at the point C where the integration over A is gaussian, but in fact the phase transition occurs on a whole "critical line" (line CT of Fig. 2), and can be interpreted as follows: inside the region PCT, where the temperature is low and the magnetic field favors spins B, the boundary of the random surface is made of spins A, whereas the bulk of the surface is mostly made of spins B, so that the boundary tries to avoid touching the rest of the surface by "collapsing on itself" (which results in a change of its Hausdorff dimension). This is a quantum gravity phenomenon which has no precise analogue on a flat lattice. Symmetrically, $a(b)$ undergoes the same phase transition on the line DT; in fact, it is clear that for any $Q \leqslant 4$, the point D exists (its position being, up to a trivial rescaling of γ, independent of Q) and is the endpoint of a boundary phase transition line, with the same critical properties as for $Q = 1$; so that we shall not mention this line in the subsequent discussion of $Q = 2, 3, 4$.

Second, let us discuss the physics on the zero magnetic field line $\alpha = \beta$, on which the Ising critical point (point T of Fig. 2) is. Note that the symmetry of the equation imposes then that $x = y$, so that there are only two unknown constants left in (3.2) (as opposed to the three found in ref. 10). A very nice picture emerges out of the patterns of the zeros of the 9th degree discriminant (Fig. 4), which in particular illustrates the $\mathbb{Z}_2$ spontaneous symmetry breaking in the low temperature phase.

At the critical point

$$\alpha_T = \beta_T = \sqrt{10} = 3.16228\ldots \qquad \gamma_T = 2\sqrt{7} + 1 = 6.29150\ldots$$

the resolvent develops a singularity ($a_\star \equiv a_2 = a_3$)

$$b - b_{\mathrm{reg}} \sim (a - a_\star)^{4/3} \tag{3.4}$$

which is characteristic of a $c=1/2$ theory coupled to gravity. Let us note that since the resolvent must have a cubic singularity, three is the minimum number of sheets required of $b(a)$; therefore the cubic two-matrix model is the simplest possible realization of the $c=1/2$ theory coupled to gravity.

Since the Ising model on random surfaces has already been studied in great detail, we shall not elaborate any further and go back to the general case with arbitrary Q. Noting that the partition function (1.5) takes the form

$$Z_Q(\alpha, \beta, \gamma) = \int dA\, e^{N \operatorname{tr}[-(\gamma/2) A^2 + (\alpha/3) A^3]} \Xi(A)^Q$$

where $\Xi(A)$ is defined by (2.1) with the usual cubic potential, and using (2.4), we can immediately write down the saddle point equations for the eigenvalues of A; after analytic continuation, we obtain:

$$(2-Q)\,\omega_A(a) + b_\star(a) + (Q-1)\,b(a) = \gamma a - \alpha a^2 \tag{3.5}$$

where $b(a) = b_+(a)$ is evaluated on the physical sheet. For $Q=2$ and 3 the resolution is very similar to the $Q=1$ case, and the resulting phase diagrams are of the same type as Fig. 2, except that the line CT becomes a real critical line (for the bulk theory); we shall now give the main results, skipping the technical details.

4. $Q=2$ AND 3 DILUTE POTTS MODELS

The matrix model corresponding to the $Q=2$ dilute Potts model on random surfaces is nothing but the $\mathbb{Z}_2$ symmetric three matrix chain,(6) A playing the role of the central matrix and B of the two matrices at the ends of the chain. If we remove the dilution, i.e., set $\alpha=0$, we are back to the Ising model without magnetic field. However, adding dilution allows us to reach the tricritical point of the Ising model.(21)

The saddle point equation (3.5) reads for $Q=2$:

$$b_\star(a) + b(a) = \gamma a - \alpha a^2 \tag{4.1}$$

which shows that $b_\star(a)$ and $b(a)$ must have the same cuts; since we know from Section 2 that $b(a) = b_+(a)$ is connected by a semi-infinite cut to $b_-(a)$, $b_\star(a) = b_{\star+}(a)$ must likewise be connected to a fourth sheet $b_{\star-}(a)$. We therefore assume that $b(a)$ satisfies a fourth degree equation and compute symmetric functions of its 4 sheets. Using once again the expansion (2.8), the saddle point equation (4.1) and various other relations coming from the analytic structure of $b(a)$, we find the following algebraic equation:

$$\beta b^4 - 2\beta(-\alpha a^2 + \gamma a)\, b^3 + (\alpha^2 \beta a^4 - 2\alpha\beta\gamma a^3 + \gamma(\beta\gamma - \alpha)\, a^2$$
$$+ (-\alpha\beta + 2 + \gamma^2)\, a + \cdots)\, b^2$$
$$- (\gamma\alpha^2 a^4 + \alpha(\alpha\beta - 2(1+\gamma^2))\, a^3 + \cdots)\, b$$
$$+ (\alpha^2 a^5 - 2\alpha\gamma a^4 + \cdots) = 0 \qquad (4.2)$$

where $\cdots$ means that there are some lower order terms whose coefficients must be fixed by the analytic structure. Note that as an equation for a, (4.2) is a fifth degree equation.

Fixing the unknown coefficients in (4.2) allows us to easily find the critical line, which is characterized by the collision of the physical cut with the semi-infinite cut. Its two endpoints are the zero dilution ($\alpha = 0$) critical point:

$$\gamma_{\mathsf{C}}^2 = 2(1+\sqrt{7}) \qquad \beta_{\mathsf{C}}^2 = 10\gamma_{\mathsf{C}}^{-3}$$

where the loop functions display the same behavior that is found along the whole critical line, namely

$$\begin{aligned} b - b_\star &\sim (a - a_\star)^{3/4} \\ a - a_\star &\sim (b - b_\star)^{4/3} \end{aligned} \qquad (4.3)$$

and are characteristic of a $c = 1/2$ theory; and the tricritical point[9]:

$$\alpha_{\mathsf{T}} = 2.83045\ldots \qquad \beta_{\mathsf{T}} = 3.09138\ldots \qquad \gamma_{\mathsf{T}} = 6.23472\ldots$$

where the corresponding singularities of the loop functions are

$$\begin{aligned} b - b_{\text{reg}} &\sim (a - a_\star)^{5/4} \\ a - a_{\text{reg}} &\sim (b - b_\star)^{5/4} \end{aligned} \qquad (4.4)$$

As expected this corresponds to the central charge $c = 7/10$ of tricritical Ising.

The 3-states dilute Potts model is the first in which the corresponding matrix model $Q = 3$ does not have the form of a linear chain; in particular, it is not solvable via orthogonal polynomials. We start once more from the saddle point equation (3.5). The discussion of the analytic structure of $b(a)$

[9] Only the numerical values are given, the exact values being fairly cumbersome. This remark also applies to the cases $Q = 3$ and 4.

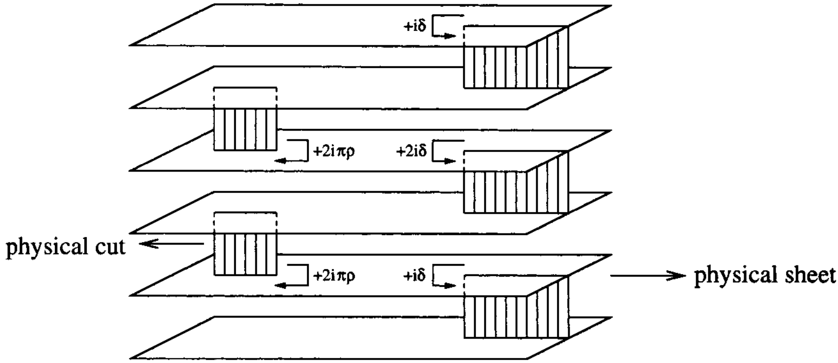

Fig. 5. Analytic structure of $b(a)$ in the 3-states dilute Potts model. Jumps of $b(a)$ across its cuts are shown.

in this case becomes a little more involved, and let us simply state the conclusion of this analysis, which is that $b(a)$ must have 6 sheets, as shown on Fig. 5. The reader can check that this analytic structure, and in particular the discontinuities shown, are compatible with the saddle point equations. Hence, we assume $b(a)$ is a solution of a sixth degree equation, and look for its coefficients as polynomials in a.

For the general dilute 3-states Potts model, the degree of these polynomials becomes quite high (up to 9), and we shall only write the algebraic equation satisfied by a and b in the non-dilute ($\alpha=0$) case:

$$\begin{aligned}
&\beta b^6+6\gamma(1-\beta a)\,b^5+(13\beta\gamma^2a^2-6(-1+4\gamma^2)\,a+\gamma(2\beta+9\gamma/\beta))\,b^4\\
&\quad+\left(-12\beta\gamma^3a^3+4\gamma(-6+7\gamma^2)\,a^2-4\frac{\gamma}{\beta}(-6+2\beta^2\gamma+3\gamma^2)\,a+\cdots\right)b^3\\
&\quad+\left(4\beta\gamma^4a^4-6\gamma^2(-5+\gamma^2)\,a^3\right.\\
&\quad\left.+\frac{1}{\beta}(9-54\gamma^2+10\beta^2\gamma^3-15\gamma^4)\,a^2+\cdots\right)b^2\\
&\quad+\left(-4\gamma^3(3+\gamma^2)\,a^4-\frac{1}{\beta}(18\gamma-24\gamma^3+4\beta^2\gamma^4-18\gamma^5)\,a^3+\cdots\right)b\\
&\quad+\left(4\gamma^4a^5+\frac{\gamma^2}{\beta}(17-18\gamma^2)\,a^4+\cdots\right)=0 \qquad (4.5)
\end{aligned}$$

Note that it was not known before that the standard resolvent of the Potts model $\omega_B(b)$ (or $a(b)$, cf Eq. (2.5)) satisfies a polynomial equation. This a fifth degree equation in a, so that $a(b)$ has five sheets. The critical point is easily determined:

$$\gamma_{\mathrm{C}}^2 = 3 + \sqrt{47} \qquad \beta_{\mathrm{C}}^2 = \tfrac{105}{4}\, \gamma_{\mathrm{C}}^{-3}$$

which is compatible with what was found in ref. 16. The singularity of the resolvents is

$$\begin{aligned} b - b_\star &\sim (a - a_\star)^{5/6} \\ a - a_\star &\sim (b - b_\star)^{6/5} \end{aligned} \tag{4.6}$$

which corresponds to a $c = 4/5$ theory.

If we introduce the dilution, the equations become rather complicated, though it is still possible to work with exact analytic expressions. We find a critical line, with singularities of the type (4.6), which ends with a tricritical point:

$$\alpha_{\mathrm{T}} = 2.44405\ldots \qquad \beta_{\mathrm{T}} = 2.9536\ldots \qquad \gamma_{\mathrm{T}} = 6.09718\ldots$$

at which the singularity becomes

$$\begin{aligned} b - b_{\mathrm{reg}} &\sim (a - a_\star)^{7/6} \\ a - a_{\mathrm{reg}} &\sim (b - b_\star)^{7/6} \end{aligned} \tag{4.7}$$

that is a $c = 6/7$ theory.

5. $Q = 4$ DILUTE POTTS MODEL

For the $Q = 4$ case, it is easy to show that $b(a)$ has an infinite number of sheets; of course, the method used above for $Q < 4$ then fails. In order to solve the model, the easiest procedure is to follow[12] in which a similar "double saddle point equation" system was solved. One introduces an auxiliary function

$$D(a) \equiv 2b(a) - \omega_A(a) - \frac{\gamma}{\beta}$$

which, for $Q = 4$ (and only $Q = 4$), satisfies a two-cut Riemann–Hilbert problem:

$$\begin{aligned} &D(a+i0) + D(a-i0) = -\alpha a^2 + \gamma a - \frac{2\gamma}{\beta} \qquad a \in [a_1, a_2] \\ &D(a+i0) + D(a-i0) = 0 \qquad a \in [a_3, +\infty] \end{aligned} \tag{5.1}$$

where $[a_1, a_2]$ is the physical cut and $[a_3, +\infty]$ is the semi-infinite cut. $D(a)$ can therefore be expressed via an elliptic parametrization in terms of Θ-functions. The critical line is expected to be found (just as in ref. 12) when the two cuts meet ($a_2 = a_3$), that is in the trigonometric limit of the elliptic functions. The solution of (5.1) becomes then much simpler, and is of the form:

$$D(a) = -\frac{\alpha}{\pi}(a-a_2)\left((a-a_0)\arctan\sqrt{\frac{a_2-a_1}{a_1-a}} + \sqrt{(a_2-a_1)(a_1-a)}\right)$$

where a_1, a_2, a_0 and the relation defining the critical line are given by the asymptotics of $D(a)$ as $a \to \infty$ and the condition $-\alpha(a-a_2)(a-a_0) = -\alpha a_2^2 + \gamma a_2 - (2\gamma/\beta)$.

It is easy to see that positivity of the density of eigenvalues implies that $a_0 \geqslant a_2$. In particular for $\alpha = 0$ ($a_0 = +\infty$), one finds

$$\gamma_{\mathrm{C}}^2 = 4(1+\sqrt{1+\pi^2/3}) \qquad \beta_{\mathrm{C}}^2 = \frac{16\pi^2}{3}\gamma_{\mathrm{C}}^{-3}$$

More generally, along the critical line one has the strict inequality $a_0 > a_2$ and the resolvent has the singularity

$$b - b_{\mathrm{reg}} \sim (a - a_\star)\log^2(a-a_\star) \tag{5.2}$$

or after inversion

$$a - a_{\mathrm{reg}} \sim \frac{b-b_\star}{\log^2(b-b_\star)} \tag{5.3}$$

On the other hand, at the tricritical point

$$\alpha_{\mathrm{T}} = 1.6523\ldots \qquad \beta_{\mathrm{T}} = 2.42087\ldots \qquad \gamma_{\mathrm{T}} = 5.4602\ldots$$

one has $a_0 = a_2$, and this reflects in a change of the leading singularity:

$$b - b_{\mathrm{reg}} \sim (a-a_\star)\log(a-a_\star) \tag{5.4}$$

or

$$a - a_{\mathrm{reg}} \sim \frac{b-b_\star}{\log(b-b_\star)} \tag{5.5}$$

The behavior (5.2) and (5.4) of the resolvent $b(a)$ is the usual one and is also found in other $c=1$ models like the $O(2)$ model;[19, 20] however, the loop functions (5.3) (in particular the standard resolvent $a(b)$ of the non-dilute 4-states Potts model) and (5.5) are more unusual, and should be compared with the similar result found in ref. 12.

6. CONCLUSION

We have discussed a new method for solving multi-matrix models in the large N limit, which relies on saddle point analysis and functional inversion relations satisfied by the unknown functions appearing in the saddle point equations. This method gives a new way of deriving equations for the resolvents of the model, and is a (much simpler) alternative to the method of loop equations for the determination of the disk amplitudes. It is not limited, like the orthogonal polynomials, to linear chains of matrices, as we have shown by solving explicitly the $(Q+1)$-matrix model representing the dilute Potts model on random surfaces.

Furthermore, we have seen that the functional inversion can be interpreted here as a duality of the theory. In the Ising model this is simply the $\mathbb{Z}_2$ symmetry of the model; but more generally; in the dilute Potts model, the functional inversion relates different boundary conditions for surfaces with the topology of a disk. On the critical line, it exchanges two different loop functions with different critical exponents ($g \to 1/g$, while the bulk theory is unchanged, cf. (1.7))—similarly to the duality that exchanges Dirichlet and Neumann boundary conditions in open string theory. At the $Q<4$ tricritical points, the universal part of the loop function turned out to be self-dual, whereas logarithmic corrections spoil the self-duality at $Q=4$.

One can hope that the ideas presented here will be applicable to various other problems of statistical mechanics on random surfaces, combinatorics, and asymptotics of orthogonal polynomials.

Note Added. After this work was completed, it was claimed in the preprint[24] that using loop equations one could reproduce the polynomial equation (4.5) of the 3-states Potts model.

APPENDIX A. CONNECTION WITH ORTHOGONAL POLYNOMIALS

We shall now show how the functions $b(a)$ and $a(b)$ can be defined via orthogonal polynomials, when such a tool is available. We shall consider the 2-matrix model only (i.e., $Q=1$), though similar results exist for an arbitrary chain of matrices (as in the $Q=2$ Potts model).

It is easy to show that the definition of $b(a)$ that was given earlier (Eq. (2.4)) is equivalent to the following:

$$b(a) = \lim_{N\to\infty} \frac{1}{N}\frac{\mathrm{d}}{\mathrm{d}a} \log Z(a) \tag{A.1}$$

with

$$\begin{aligned} Z(a) = \iint \prod_{i=1}^{N} \mathrm{d}a_i\, \mathrm{e}^{-NW(a_i)} \prod_{i=1}^{N+1} \mathrm{d}b_i\, \mathrm{e}^{-NV(b_i)} \Delta_{i=1\cdots N}[a_i] \\ \times \Delta_{i=1\cdots N+1}[b_i] \det_{i,j=1\cdots N+1}[\mathrm{e}^{Na_ib_j}]_{a_{N+1}\equiv a} \end{aligned} \tag{A.2}$$

where V and W are the polynomial potentials. If we now introduce bi-orthogonal polynomials $P_n(a) = a^n + \cdots$ and $Q_m(b) = b^m + \cdots$ such that

$$\iint \mathrm{d}a\, \mathrm{d}b\, P_n(a)\, Q_m(b)\, \mathrm{e}^{N(ab - W(a) - V(b))} = \delta_{mn} h_m \tag{A.3}$$

then after reordering the a_i we find that

$$\begin{aligned} Z(a) = N! \iint \prod_{i=1}^{N} \mathrm{d}a_i\, \mathrm{e}^{-NW(a_i)} \prod_{i=1}^{N+1} \mathrm{d}b_i\, \mathrm{e}^{-NV(b_i)} \\ \times \det_{i,n=1\cdots N}[P_n(a_i)] \det_{j,m=1\cdots N+1}[Q_m(b_j)]\, \mathrm{e}^{N(ab_{N+1} + \sum_{i=1}^{N} a_ib_i)} \\ = N! \prod_{i=1}^{N} h_i \int \mathrm{d}b\, Q_N(b)\, \mathrm{e}^{Nab}\, \mathrm{e}^{-NV(b)} \end{aligned} \tag{A.4}$$

and therefore

$$b(a) = \frac{\int \mathrm{d}b\, bQ_N(b)\, \mathrm{e}^{Nab}\, \mathrm{e}^{-NV(b)}}{\int \mathrm{d}b\, Q_N(b)\, \mathrm{e}^{Nab}\, \mathrm{e}^{-NV(b)}} \tag{A.5}$$

which shows that b is directly related to $\hat{b}$, the operator which acts on orthogonal polynomials as multiplication by b. If we now introduce the shift operator $\hat{s}$ such that

$$\hat{s}Q_n = c_n^{-1} Q_{n+1} \tag{A.6}$$

with the appropriate normalization faction $c_n = \sqrt{h_n/h_{n+1}}$, then we know that acting on the Q_n with $N - n \ll N$,

$$\hat{b} = c\hat{s} + \sum_{k=0}^{\deg V - 1} b_k \hat{s}^{-k} \tag{A.7}$$

($c \equiv c_N$) and similarly by making $\hat{s}$ act on the P_m, the operator $\hat{a}$ of multiplication by a takes the form

$$\hat{a} = c\hat{s}^{-1} + \sum_{k=0}^{\deg W-1} a_k \hat{s}^k \tag{A.8}$$

Returning to Eq. (A.5), we conclude that

$$b(a) = s(a) + \sum_{k=0}^{\deg V-1} b_k s(a)^{-k} \tag{A.9}$$

where $s(a)$ is the eigenvalue of $\hat{s}$ corresponding to the eigenvalue a of $\hat{a}$, and similarly for $a(b)$.

For example, in the case of a cubic potential, what we have found is the following parametrization of a and b:

$$\begin{cases} b = cs + b_0 + b_1 s^{-1} + b_2 s^{-2} \\ a = cs^{-1} + a_0 + a_1 s + a_2 s^2 \end{cases} \tag{A.10}$$

The interesting property of this parametrization is that it automatically implies the analytic structure found in Section 3. In particular, it gives us an alternative way to fix the constants x, y, z of Eq. (3.2).

ACKNOWLEDGMENTS

I thank I. Kostov for e-mail discussions, and D. Johnston and M. Stathakopoulos for pointing out ref. 23 to me. This work was supported in part by the DOE Grant DE-FG02-96ER40559.

REFERENCES

1. E. Brézin, C. Itzykson, G. Parisi, and J.-B. Zuber, *Commun. Math. Phys.* **59**:35 (1978).
2. V. A. Kazakov, *Mod. Phys. Lett. A* **4**:2125 (1989).
3. C. Itzykson and J.-B. Zuber, *J. Math. Phys.* **21**:411 (1980).
4. H. Chandra, *Amer. J. Math.* **79**:87 (1957).
5. M. L. Mehta, *Comm. Math. Phys.* **79**:327 (1981).
6. S. Chadha, G. Mahoux, and M. L. Mehta, *J. Phys. A* **14**:579 (1981).
7. D. V. Boulatov and V. A. Kazakov, *Phys. Lett. B* **186**:379 (1987).
8. J. Alfaro, *Phys. Rev. D* **47**:4714 (1993).
9. D. V. Boulatov, *Mod. Phys. Lett. A* **8**:557 (1993).
10. M. Staudacher, *Phys. Lett. B* **305**:332(1993).
11. P. Zinn-Justin, *Commun. Math. Phys.* **194**:631 (1998).
12. V. A. Kazakov and P. Zinn-Justin, to be published in *Nucl. Phys. B.*
13. E. Brézin and D. J. Gross, *Phys. Lett. B* **97**:120 (1980).

14. I. Kostov, *Nucl. Phys. B (Proc. Suppl.)* **10A**:295 (1989); D. J. Gross and M. J. Newman, *Phys. Lett. B* **266**:291 (1991).
15. V. A. Kazakov, *Nucl. Phys. B (Proc. Suppl.)* **4**:93 (1988).
16. J. M. Daul, preprint hep-th/9502014.
17. V. A. Kazakov, M. Staudacher, and T. Wynter, *Commun. Math. Phys.* **179**:235 (1996).
18. G. Moore, N. Seiberg, and M. Staudacher, *Nucl. Phys. B* **362**:665 (1991).
19. I. Kostov, *Nucl. Phys. B* **376**:539 (1992).
20. I. Kostov and M. Staudacher, *Nucl. Phys. B* **384**:459 (1992).
21. K. Fukazawa, K. J. Hamada, and H. Sato, *Mod. Phys. Lett. A* **5–29**:2431 (1990).
22. B. Nienhuis, A. N. Berker, E. K. Riedel, and M. Schick, *Phys. Rev. Lett.* **43**:737 (1979).
23. V. A. Kazakov, *Mod. Phys. Lett. A* **4**:1691 (1989).
24. G. Bonnet, preprint hep-th/9904058.