An essay on the dynamics and statistics of critical field theories*

by

Robert P. Langlands

1. Introduction. This article is even more provisional and premature than its title suggests. Its author is entitled to no pretensions. Field theories and especially conformally invariant field theories are becoming familiar to mathematicians, largely because of their influence on the study of Lie algebras and above all on topology. Nonetheless, in spite of the progress in constructive quantum field theory during recent decades, many analytic problems, especially the existence of the scaling limit, are given short shrift. These problems are difficult and fascinating and merit more attention. Some important basic concepts, due to physicists, are available but, so far as I can see, need substantial mathematical elaboration.

It is often overlooked that the largely *mathematical* development of Newtonian mechanics in the eighteenth century was an essential prerequisite to the enormous *physical* advances of the nineteenth and twentieth centuries, that attempts to overcome mathematical obstacles may lead to concepts of physical significance, and that mathematicians, recalling the names of d'Alembert, Lagrange, Hamilton and others, may aspire to nobler tasks than those currently allotted to them. Age and limited talent and knowledge foreclose all such ambitions to me; indeed for lack of experience my intuitions as to the real issues may be quite off base. Nonetheless the informal forum provided by the present collection is a welcome occasion for self-indulgence.

My goal is modest. After recalling, in a laconic fashion, some problems posed by scaling limits for the Ising model, I review my – partly collaborative – attempts to introduce in the simpler context of percolation an ultraviolet cutoff in closed form. The constructions utilised for percolation suggest, in a rough way, analogous constructions for models with interaction. These constructions are examined not for the Ising model but for free particles, and lead to a rather elegant paradigm that I hope will also, with suitable modifications, be applicable to the Ising model. Since obscurities and uncertainties remain, it would be premature to say more, for the basic definitions for the Ising model cannot be made without a better understanding of several aspects of it than we appear to have at present. The reflections on free particles, especially the calculations in Section 6 establishing the existence of the map ξ, may be of some interest in themselves, even though the ultimate purpose is to suggest appropriate and useful questions about the Ising model.

My pedestrian efforts have been encouraged, sometimes unwittingly and occasionally to their regret, by several colleagues and friends whom I would like to thank explicitly, even though they might prefer to distance themselves from the endeavour. The focus of any attempt to deal with the existence of the scaling limit for all but artificially simple cases is necessarily the Ising model ([MW]). Although the very deepest problems appear to be as difficult for it as for all other models, extremely important work by Onsager and others yield specific results that guide all reflection and, in particular, allow a prompt rejection of many impetuous hypotheses. Many of the observations to follow arose from the continuing attempt, in collaboration with Marc-André Lewis and Yvan Saint-Aubin, to acquire a basic understanding of the dynamics of renormalization of the Ising model. My first thanks go to them, and to Marc-André Lafortune. My second thanks go to Oliver Sick and Axel Schmitz-Tewes who with great forbearance and loyalty attended a course at the University in Bonn sponsored by the Max-Planck-Institut in which some of the half-baked ideas presented here were discussed in raw form, and to Günter Harder and Friedrich Hirzebruch for the invitation to deliver the course. Specific thanks are due to Burton Randol for comments on the Euler-Maclaurin formula, to Fan Chung who explained the identity of Saalschütz and how to use it in the present context, and to John Palmer for comments on the literature. I am also obliged to William Casselman, Daniel Friedan, Paul Phong, Thomas Spencer, Gerard Watts, and Jan Wehr whose observations have at critical junctures guided my thoughts.

* First appeared in Can. Math. Soc., Selecta 1945-1995, Vol. 3, 1996. Included by permission of Canadian Mathematical Society.

2. The Ising model. The passage from lattice models to field theories is, in principle, carried out in two steps ([GJ]): the first from lattice models to *euclidean* field theories; the second from euclidean field theories to *minkowskian* field theories. For various reasons the second step is often omitted so that a field theory is often taken to be a euclidean theory. Its difficulties are certainly of a much different nature than those attached to the first step, which is the existence of the scaling limit and around which this article turns,

I review the problems in the context of the Ising model. Consider the planar lattice $\Lambda = \mathbb{Z}^2$ and the subset $\Lambda_N = \{(x,y) \mid |x| \le N,\, |y| \le N\}$. If s is a function on Λ_N with values in $\{\pm 1\}$ set

$$H(s) = H(s,K) = -\sum_{p,q} K s(p)s(q).$$

The sum runs over nearest neighbours p and q in L_N and K is a positive parameter, whose inverse T corresponds to the temperature. (Mathematically T is superfluous, but not to use it leads inevitably to inequalities with the false sense!) The Boltzmann factor attached to s is $\exp(-H(s))$; it defines a probability on the space of all such s, the probability of t being

$$\frac{\exp(-H(t))}{\sum_s \exp(-H(s))}. \tag{2.1}$$

The denominator, in which the sum runs over all s, is called the partition function and denoted $Z = Z(K,N)$.

The probability space defined, various expectations can be introduced, of which the most usual are the correlation functions

$$E(s(p_1)\dots s(p_r)) = E(s(p_1)\dots s(p_r); K, N). \tag{2.2}$$

Here the p_i are points in Λ_N so that each $s \to s(p_i)$ is a function on the probability space as is their product. It is the expected value of the product that appears in (2.2). For a given K and arbitrary points p_i, $1 \le i \le r$ the limit

$$\lim_{N\to\infty} E(s(p_1)\dots s(p_r); K, N) = E(s(p_1)\dots s(p_r); K) \tag{2.3}$$

exists. These limits are expectations for a probability measure, referred to as a Gibbs state, on the space of functions on Λ with values in $\{\pm 1\}$.

A further important possibility is to consider only those s whose restrictions to the boundary of Λ_N are constrained to be equal to a given function σ. The sum $Z(K,N;\sigma)$ in the denominator of (2.1) is then taken over the collection of such functions, and the expectations in (2.2) are replaced by

$$E(s(p_1)\dots s(p_r); K, N; \sigma).$$

The sequences appearing in (2.3) are then associated to a sequence $\{\sigma_N\}$ of boundary values and only for appropriate choices of these boundary values will the limits exist for a given K and arbitrary p_i, again defining a Gibbs state.

There is a critical value $K_c = 0.440687\dots$ defined by $\sinh 2K_c = 1$; the associated value $T_c = 1/K_c$ is the Curie point. For T greater than or equal to the critical value T_c the limits that exist are all the same, and equal to those obtained directly from (2.3), so that there is only one Gibbs state. For T less than this value they are not, but the manifold of possibilities is quite well understood ([MMS]). In particular there are only two translation-invariant extremal measures that are then characterised by the sign of the magnetisation, namely the expectation of $s(p)$, p arbitrary. For $T < T_c$ the two extreme Gibbs states yield two possibilities for the correlations that differ only by a sign,

$$E_+(s(p_1)\dots s(p_r); K) = (-1)^r E_-(s(p_1)\dots s(p_r); K). \tag{2.4}$$

The Gibbs state defined by (2.3) is then the mean of these two states. For the sake of a uniform notation we let

$$E_+(s(p_1)\dots s(p_r); K) = E_-(s(p_1)\dots s(p_r); K) \tag{2.5}$$

be the correlation functions defined by (2.3) if $K \le K_c$. This equation is compatible with (2.4).

Although for these definitions the lattice Λ_N is fixed, we could imagine it as provided with a mesh α that, for example, we could take smaller and smaller obtaining a finer and finer grid on the plane. Thus the correlation functions are defined on different subsets of the plane. It is best simply to define

$$\epsilon_\pm(\alpha p_1, \ldots, \alpha p_r; K; \alpha) = \lambda^{-r}(\alpha) E_\pm(s(p_1) \ldots s(p_r); K) \tag{2.6}$$

where the correlation functions on the right are those of (2.4) and (2.5) and $\lambda(\alpha)$ is a factor whose purposes are still to be explained. The relation

$$\lambda^r(\alpha)\epsilon_\pm(\alpha p_1, \ldots, \alpha p_r; K; \alpha) = \lambda^r(\alpha')\epsilon_\pm(\alpha' p_1, \ldots, \alpha' p_r; K; \alpha') \tag{2.7}$$

is immediate.

The correlation functions of a euclidean field theory are then obtained as limits as the mesh approaches 0. The critical value K_c plays a particular role in the construction of these limits because K must approach K_c as α approaches 0. Define, for example, $K_\alpha^\pm$ as the solution of the equation

$$\alpha = \frac{|\tanh^2 K + 2\tanh K - 1|}{\tanh K(1 - \tanh^2 K)},$$

for which $\pm(K - K_c) > 0$, so that $K_\alpha^\pm - K_c = \pm(\sqrt{2} + 1)\alpha + \mathrm{O}(\alpha^2)$. Observe that $\sinh 2K_\alpha^+ \sinh 2K_\alpha^- = 1$. If $x \ge 0$ set $K_\alpha^\pm(x) = K_c + x(K_\alpha^\pm - K_c)$. Then the limits to examine are

$$\lim_{\alpha \to 0} \epsilon_\pm(p_1, \ldots, p_r; K_\alpha(x); \alpha) = S_\pm(p_1, \ldots, p_r; x). \tag{2.8}$$

The issue on which the reflections of this paper turn is that the difficulty of rigorously establishing the existence of these limits for $x = 0$ seems to be of an entirely different order of magnitude than for $x > 0$. (If one likes, the difference is between a massless theory and a theory with positive mass.) Even for $x > 0$ it is a very serious problem that has been solved by Palmer and Tracy ([PT1], but see also [SO]) whose treatment is based on a long sequence of important contributions that began with Onsager's famous paper ([O]). There is a helpful historical review of the development in the introduction to their paper.* The factor $\lambda(\alpha)$ is defined in their paper. It is

$$(1 - \sinh^{\mp 4}(2K_\alpha^\pm))^{1/8}.$$

It approaches 0 as $K_\alpha \to K_c$. All that really matters is that $\lambda(\alpha) \sim \alpha^{1/8}$.

A glance at the definition (2.6) suggests that limits (2.8) have to be taken in a special sense: the points p_i must have rational coordinates and α goes to 0 through the partially order set $\{1/n | n \in \mathbb{N}\}$, where $1/m$ is taken to be greater than $1/n$ if n divides m. To define S everywhere a supplementary argument is then necessary. The limits are in fact only taken at collections of points $\{p_i | 1 \le i \le r\}$ all of whose coordinates are different, but with this reservation Palmer and Tracy observe that their formulas give a meaning to (2.6) for all α and all points so that this detour is unnecessary.

Like that of $\lambda(\alpha)$, the precise definition of $K_\alpha^\pm$ is of little importance. If $x \ne 0$ we can define $\alpha = \alpha(\beta, x)$ as a function of β and x by the condition that $K_\alpha^\pm(x) - K_c = \beta$ and then take the limit of (2.8) as $\beta \to 0$. Observe that $\alpha(\beta, x)/\alpha(\beta, y) \to y/x$ and that $\lambda(\alpha(\beta, x))/\lambda(\alpha(\beta, y)) \to (y/x)^{1/8}$. If $a > 0$ and $y = ax$ then, granted the appropriate uniform convergence (see p. 372 of [PT1]), it follows immediately from (2.7) that

$$\frac{S_\pm(ap_1, \ldots, ap_r; x)}{S_\pm(p_1, \ldots, p_r; y)} = \lim_{\beta \to 0} \frac{\epsilon_\pm(a(\beta)p_1, \ldots, a(\beta)p_r, \beta + K_c, \alpha(\beta, x))}{\epsilon_\pm(p_1, \ldots, p_r, \beta + K_c, \alpha(\beta, y))} = a^{-r/8},$$

* The papers of Palmer-Tracy and of Schor-O'Carroll ([SO]) are technically elaborate. I do not pretend to have studied either with the care they deserve, but there appears to be very widespread misunderstanding in the mathematical physics community of the nature of the problems associated with the existence of the scaling limit and of the extent to which they have been treated. The contributions of these two pairs of authors permit a clear understanding of the limits of present methods.

with $a(\beta) = \alpha(\beta, x)/\alpha(\beta, y)$ so that

$$S_{\pm}(ap_1, \ldots, ap_r; x) = a^{-r/8} S_{\pm}(p_1, \ldots, p_r; y). \tag{2.9}$$

It is of course no surprise that the strength K of the interaction changes with the mesh. What is intuitively appropriate is that the effective strength of the interaction between sites at a fixed distance remain approximately constant – whatever this might mean. The analysis fixes a meaning: the interaction between nearest neighbours on the fine lattice differs from K_c by a term of the order of the mesh. This is because at K_c (or T_c) the balance between, on the one hand, the multitude of paths effecting the propagation from neighbour to neighbour and, on the other, the strength of the interaction (in contrast to the imbalance at $K < K_c$ where the strength of the interaction is too weak) yields a residual propagation through arbitrarily large distances that (now in contrast to the effect of the strong interaction for $K > K_c$) does not entail any conformity in the spins. The balance is delicate and rigorous treatments of its consequences are not available.

The hypothesis, which admits no real doubt and for which a tremendous amount of evidence of various sorts is available, is that for any integral points p_i, $1 \leq i \leq r$,

$$\lim_{m\to\infty} m^{r/8} E(s(mp_1)\ldots s(mp_r)); K_c) = S(p_1, \ldots, p_r) \tag{2.10}$$

exists provided that there are no coincidences among the p_i. The same formula permits their definition at rational points and then, if the appropriate continuity is established, at all points. At $K = K_c$ the subscript $\pm$ is unnecessary. The expression in (2.10) is, in view of (2.6), essentially (2.8) at $x = 0$. Thus, failing a direct attack on the limits of (2.10), one can attempt to define $S(p_1, \ldots, p_r)$ as

$$\lim_{x\to 0} S_{\pm}(p_1, \ldots, p_r; x), \tag{2.11}$$

although the indeterminacy in λ entails an indeterminacy in (2.11) that is not present in (2.10). To what extent these alternative limits are known to exist is not yet clear to me. In view of (2.9) their existence is also an assertion about the asymptotic behaviour in $p_1, \ldots, p_r$ (for small separations and fixed x) of the functions $S_{\pm}(p_1, \ldots, p_r; x)$ about which there is something known but much less than expected. (See the discussion of the literature in [PT2].) For $r = 2$ the existence is a simple consequence of formula 2.32 of [WMTB], but this formula is no easy matter. (See [T] in which its proof is completed.) Even the problems arising from an indirect attack, in which there are many fascinating fine points, have not had all the attention they deserve.

A direct treatment of the limits (2.10) is nonetheless to be preferred, especially a treatment that is not based on a detailed understanding of the algebraic structure specific to the Ising model, but that is much more robust – compatible with perturbations and applicable to critical points of much different models.

A very general and very powerful concept is available, that of renormalization. For planar lattice models such as the Ising model its intuitive content is very simple. We are interested in the statistical properties of the models on very large finite lattices Λ_N. They are the same for all large N. If we take four such lattices we can paste them together to form the lattice Λ_{2N}. Since $2N$ is, like N, just a large integer, this operation can be interpreted as a dynamical system in which the basic map transforms one model into another. The critical points, at which the scaling limits can be formed without varying the strengths or the nature of the interactions, can be interpreted as fixed points of the map. The difficulty with this notion, blatantly scamped in the preceding remarks, is one of closure: $2N$ is not N.

3. Percolation. The hope that informs this article, and several other, usually collaborative, efforts of its author is that the problem of closure can be treated by finite-dimensional approximations to the renormalization-group transformations. I review the results for percolation ([LPPS,LPS,LL]).

The elements are: a space, thus a collection of coordinates; a collection, with index ι, of finite-dimensional approximations to it, thus finite collections of these coordinates; and for each finite-dimensional approximation an explicit transformation Θ_ι. There is to be moreover: a point η in the space *naturally* associated to the critical lattice model; for each ι a fixed point ν_ι of Θ_ι such that the collection $\{\nu_\iota\}$ approximates η and the critical indices of Θ_ι at ι approximate the critical indices attached to the critical model.

For two-dimensional percolation such objects were introduced in [LPS] and [LL]. The coordinates are indexed by simple, closed, smooth curves C in the plane and by arcs $\alpha_1, \ldots, \alpha_m$, $\beta_1 \ldots, \beta_m$, $\gamma_1, \ldots, \gamma_n$, $\delta_1, \ldots, \delta_n$ of C. Thus there is an abundance of coordinates, many of which turn out to be redundant. At present, the association of a point in the space defined by these coordinates to a lattice model is a procedure that is possible experimentally but that presupposes the existence of limits [LPPS]. These limits have not been shown to exist. Moreover an important aspect of *universality* is that the point obtained is *largely* independent of the particular critical model of percolation chosen. It is a question of symmetry; here we implicitly consider only models that are symmetric with respect to the interchange of the two axes and to a reversal of the orientation of one or the other axis, so that the adverb *largely* is unnecessary. This said I recall briefly the assignment of the coordinate

$$\pi(C, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n, \delta_1, \ldots, \delta_n) \tag{3.1}$$

to site percolation on the square lattice at the critical probability $p = p_c = .5927439\ldots$ (The assignment is still possible for $p > p_c$ or $p < p_c$ but the point obtained is uninteresting.)

Recall that the sites are open with probability p, and for a given state an open path is a path that passes from one point to a nearest neighbour, all points on the path being open. If $A > 0$ is a large number then we may dilate C and the given arcs on it by A and ask for the probability

$$\mathrm{prob}(C, \alpha_1, \ldots, \alpha_m, \beta_1, \ldots, \beta_m, \gamma_1, \ldots, \gamma_n, \delta_1, \ldots, \delta_n; A) \tag{3.2}$$

that there is for each i an open path within C joining α_i to β_i but for each j no open path joining γ_j to δ_j. This notion is somewhat approximate but the limit of (3.2) as $A \to \infty$ should have a meaning and this is (3.1). The interest of these coordinates can be regarded as well-established.

It turns out – experimentally – that it is enough to know those attached to a square C of side 1 (See [LPS].) Moreover by a presumed continuity it is enough to know, for each positive integer l, the probabilities of events defined by a subdivision of each side of the square into subintervals of length $1/l$. More precisely, fixing l, consider a function ϕ from pairs $\{\alpha, \beta\}$ of the $4l$ intervals so obtained with values in $\{0, 1\}$. Then consider as before when the square and the intervals are dilated by a factor A the probability that two dilated intervals α and β are joined by an open path in the interior of the dilated square if and only $\phi(\alpha, \beta) = 1$. The limits of these probabilities yield a function π_l on the collection of functions ϕ.

The finite-dimensional approximating spaces are indexed by positive integers l, supposed multiplicatively ordered, and these approximating spaces are collections Π_l of measures on the set $\mathfrak{Y}_l$ of functions ϕ on pairs of intervals of length $1/l$. There are constraints on the functions and the measures as well as some slight technical modifications in the definitions, that are all explained fully in [LL]. They need not concern us here. These spaces form a projective system: there is an obvious map from Π_l to Π_k if $k|l$. There is also a map from the full infinite-dimensional space to each of the finite-dimensional spaces Π_l. Let η_l be the image in Π_l of the point η attached to critical percolation and defined by the coordinates (3.1). In essence it is the same as π_l.

The map Θ_l is defined in [LL]. I do not repeat all details of the definition here. A point in $\mathfrak{Y}_l$ is just a possible and coherent response to the collection of questions whether the interval α is joined to β for the $16l^2$ possible pairs of intervals of length $1/l$. Put four squares together to form a larger square, and choose four points $\phi_{i,j}$, $1 \le i, j \le 2$ in $\mathfrak{Y}_l$, one attached to each of the squares. Then try to pass from an interval on a side of the larger square to another such interval using the joins in the smaller squares given by these four points in $\mathfrak{Y}_l$. This defines a map from $\mathfrak{Y}_l \times \mathfrak{Y}_l \times \mathfrak{Y}_l \times \mathfrak{Y}_l$ to $\mathfrak{Y}_l$ and thus a map from Π_l to Π_l. This map is Θ_l.

It is easy to establish the existence of an interesting fixed point ν_1 of Θ_1. Moreover the purpose of [LL] was the numerical investigation of Θ_2. In particular it was shown that ν_1 could be lifted (approximately!) to a fixed point ν_2 of Θ_2 that is also a reasonable approximation to η_2. The analytic problem, difficult and not yet seriously broached, is to continue these approximate liftings through a sequence of larger and larger l to arrive at approximations to η. Since the unstable directions at the fixed point appears already at level one, it is only a question, as in Lanford's treatment of the Feigenbaum cascade, of adding stable directions and this is, in principle, feasible. The difficulty is that, in contrast to maps of the interval, for percolation the source of the stability remains unclear.

Moreover, it is not for percolation alone that the difficulties of dealing with universality are unresolved. The problem is there, in even more acute forms, for other lattice models, and is intimately associated with that of the existence of the scaling limit at criticality. Thus the introduction of finite models, as in [LL], has to be justified in two quite different respects, both important: the possibility of repeatedly refining the approximations has to be established; the necessary constructions have to be extended to other models, at least to the Ising model, which remains the touchstone. We shall be concerned in the following pages with the second problem. For a view of the critical point in percolation not unrelated to that of the present section, see [A].

4. Partition functions. The intuition that informs the paper [C3], and that is shared I suppose by all investigators, is that the crossing probabilities (3.1) are degenerate forms of partition functions. So partition functions are likely to appear in the finite-dimensional dynamical systems for nondegenerate models, thus for those in which there is an interaction between different sites.

To recall briefly in the pertinent context the notion of a partition function, I consider statistical systems on a lattice in a rectangle R of sides A and B and of mesh $1/L$ so that there are $(A+1)(B+1)L^2$ sites in the rectangle and $2A+2B+4L$ sites on the boundary. (The business of counting points in the rectangle and on the boundary is somewhat rough. A slight shift in the position of the rectangle in relation to the lattice can cause a change of order L in the number of sites in the rectangle and of order 1 in the number of sites on the boundary. This clearly wreaks havoc with the behaviour suggested below. An adequate formulation would require the introduction of averages over L, but this is at the moment a fine point.) At each site there is some system (described by points in a finite set or in $\mathbb{R}$) that can interact with the systems at neighboring sites, so that if the systems are prescribed at all sites by a function s (with values in the appropriate set) then there is attached to s an energy $H(s)$. We fix some function σ on the boundary ∂R and define the value of the *partition function* at σ to be

$$Z(\sigma) = \sum_{s|\partial R=\sigma} e^{-H(s)}.$$

For example in the Ising model, the set describing the system at a site is $\{\pm 1\}$ and

$$H(s) = -\sum K s(p)s(q),$$

where the sum is over all nearest neighbors inside R or on its boundary and K a positive constant. A typical choice of σ would be a the constant $+1$ or the constant -1. We could also take it to be $+1$ on the vertical sides and -1 on the horizontal sides. It is to be stressed that, in this example and in general, σ is not to depend on the mesh of the lattice.

The behaviour of $Z(\sigma)$ for fixed σ but for $L \to \infty$ is generally expected to be

$$Z(\sigma) = \exp(a\alpha L^2 + b\lambda L + d\ln L + e), \tag{4.1}$$

where a, the free energy per site, is a constant that does not depend on σ as is (I believe) d. The area enclosed by the rectangle is α and its perimeter is λ. The constant b may or may not depend on σ. At the moment I prefer not to be specific about b, not being too sure as to its possible form. The important term for my purposes is e because it seems reasonable to expect that

$$\lim_{L\to\infty} \exp(e) = z(\sigma) \tag{4.2}$$

exists. I shall sometimes refer to $z(\sigma)$ as the reduced partition function. It is also usually referred to as the partition function, the distinction made here being overlooked.

Such behaviour is expected not only for rectangular regions, but for general planar regions, and even, for appropriately defined models, for regions on two-dimensional Riemannian manifolds. (Indeed the problem posed initially is for models in any dimension. The partition function $Z(\sigma)$ can still be defined and its behaviour is presumably such that $z(\sigma)$ too exists. The principal reason for considering two-dimensional models is that for them conformal invariance, expected to be valid for the scaling limits in all dimensions, yields a richer structure and thus more formulas.) Some evidence will be presented below. There is more.

The function $\sigma \to z(\sigma)$ can also be regarded, after normalization (of some sort, for we are not excluding the possibility that there are an infinite number of possibilities for σ), as a measure on the set of all σ. As for percolation we would like to find an increasing sequence of sets Σ_l, $l = 1, 2, 3, \ldots$, each a collection of possible boundary values for the σ and each finite (or finite-dimensional), and functions z_l (or measures – according to the point of view) on Σ_l such that z_l approximates z and is a fixed point of an appropriately defined *renormalization* transformation Θ_l. For the Ising model itself, a difficulty is quickly encountered: the appropriate boundary conditions are certainly not those one would naively expect. For this and other reasons, there are better models to examine at this stage.

The conformally invariant theories for free bosons are closely related to Laplace's equation, and they are particularly easy to examine closely.

I recall the basic theory from [G]. The boundary conditions σ are of the form $\exp(i\varphi/\rho)$ where ρ is a parameter of the theory and φ a real-valued function. Thus the function s on sites is $s(p) = \exp(if(p)/\rho)$, where f is again a real-valued function. The energy is

$$H(f) = \frac{1}{2\pi}\sum_{p,q}(f(p) - f(q))^2, \tag{4.3}$$

where the sum is over nearest neighbors.

The boundary of the *open* region R under consideration is supposed to be a finite collection of simple, closed, smooth curves and the region to be bounded. The boundary conditions are defined by a function φ that is smooth in a neighbourhood of the boundary curve. If Λ is a fine square lattice (containing for the sake of precision the origin) of mesh $1/L$ then the set Λ_R of lattice points in R is well defined as is the boundary $\partial\Lambda_R$ of Λ_R. It is the collection of points in Λ_R with an exterior point as nearest neighbour. Thus, in principle,

$$Z(\sigma, R) = Z(\sigma) = Z(\varphi) = \sum_{\varphi'}\int e^{-H(f)} = \sum Z_{\mathrm{part}}(\varphi').$$

The sum is taken over all functions φ' on the boundary such that the value of $\varphi' - \varphi$ at each point of the boundary lies in $2\pi\mathbb{Z}$, but modulo the relation $\varphi_1' \sim \varphi_2'$ if $\varphi_1' - \varphi_2'$ is constant; the integral, a Gaussian integral, is taken over the affine space of all functions f on Λ_R whose restriction to the boundary is equal to φ'. (Observe that a Gaussian integral over a real affine space involves two forms, one positive – defining the Lebesgue measure on the space – and one non-negative – defining the Gaussian factor. The Lebesgue measure is defined by $\sum f(p)^2$, the sum being over all interior points.)

Let $D_\Lambda(\varphi')$ be $H(\tilde\varphi')$ where $\tilde\varphi'$ is the function on Λ_R equal to φ' on the boundary and harmonic (in the discrete sense) in the interior. Then $Z(\varphi')$ equals

$$\sqrt{\pi}^{N(\Lambda_R)}\Delta^{-1/2}\exp(-D_\Lambda(\varphi')) \tag{4.4}$$

Where $N(\Lambda_R)$ is the number of points in the interior of Λ_R and Δ is the product of the eigenvalues of the difference operator

$$-\{f(x+1, y) + f(x-1, y) + f(x, y+1) + f(x, y-1) - 4f(x)\}/2\pi$$

on Λ_R with zero boundary values.

If the domain R that we are treating is the square $0 < x, y < 1$ of side 1 and the mesh is $1/L$, L integral, then $N(\Lambda_R) = (L-1)^2$ and

$$\Delta = \prod_{k,l=1}^{L-1} 2(\sin^2(k\pi/2L) + \sin^2(l\pi/2L))/\pi.$$

The asymptotic behaviour for large L of $\Delta^{-1/2}$ and thus of the factor in (4.4) can be found with a two-dimensional form of the Euler-Maclaurin sum formula. The result is*

$$\pi^{(L-1)^2}\Delta^{-1/2} \sim 2^{-(L-1)^2/2}\pi^{(L-1)^2}\exp(\alpha L^2+\beta L+\gamma\ln L+\epsilon) \sim \exp(\alpha' L^2+\beta' L+\gamma\ln L+\epsilon')$$

with

$$\alpha = -\frac{1}{2}\int_0^1\int_0^1 \ln(\sin^2(\pi x/2)+\sin^2(\pi y/2))dxdy = .110025, \qquad \alpha' = \alpha - \ln 2/2 + \ln\pi = 4.90191$$

$$\beta = \frac{1}{2}\int_{\partial R}\ln(\sin^2(\pi x/2)+\sin^2(\pi y/2)) = -0.50492, \qquad \beta' = \beta + \ln 2 + 2\ln\pi = 2.93212$$

$$\gamma = 1/4 \qquad .$$

I did not calculate ϵ. The behaviour is that predicted in formulas (4.1), (4.2), and (4.3); and

$$z(\sigma) = z(\varphi) = \lim_{L\to\infty}\exp(\epsilon)\sum_{\varphi'}\exp(-D_\Lambda(\varphi')).$$

If the values of φ' change sharply then $D_\Lambda(\varphi')$ is large; and it is to be supposed for present purposes, although I have not tried to verify it, not only that the terms associated to a φ' for which $\varphi'-\varphi$ is not a constant $2\pi m$ approach 0 but also that their sum does, so that for a region with connected boundary

$$z(\sigma) = \lim\exp(\epsilon)\exp(-D_\Lambda(\varphi)).$$

In [C2] Cardy predicts on the basis of "physical" considerations the value

$$4\frac{c\vartheta}{24\pi}((\pi/\vartheta)^2-1).$$

for $-\gamma$. Since the central charge c is for free bosons equal to 1 and ϑ is the interior angle at a corner of the square, thus $\pi/2$, this expression is $-1/4$. The agreement is welcome. There appear to be few rigorous studies in the mathematical literature of the asymptotic behaviour of partition functions (even for simple planar domains and even for free fields) as the mesh of the lattice goes to 0.†

For free bosons, there is an obvious passage to the limit $L\to\infty$, although its implications are not all manifest. The function φ becomes a continuously differentiable function on the boundary of the square (or of any other region R being treated). Such a function can be extended to a function $\tilde\varphi$ harmonic in the interior, and we can introduce

$$\begin{aligned} D(\varphi) &= \frac{1}{2\pi}\int_R\Big((\frac{\partial\tilde\varphi}{\partial x})^2+(\frac{\partial\tilde\varphi}{\partial y})^2\Big)dxdy \\ &= \frac{2}{\pi}\int_R\frac{\partial\tilde\varphi}{\partial z}\frac{\partial\tilde\varphi}{\partial\bar z}dxdy \end{aligned}$$

and expect (although I have not examined the pertinent literature on numerical analysis) that

$$z(\varphi) = \exp(-D(\varphi)+\epsilon).$$

It is also possible to calculate formally in the limit itself. Then

$$Z(\varphi) = \int_R\exp(-H(f)), \qquad H(f) = \frac{1}{2\pi}\int_R\Big((\frac{\partial f}{\partial x})^2+(\frac{\partial f}{\partial y})^2\Big)dxdy,$$

* Fan Chung observed, on correcting my original calculation, that there are more efficient and combinatorially more interesting ways to make it.

† Oliver Sick has drawn my attention to [DD] and to [F].

the integral being a formal Gaussian integral over the set of function whose restriction to the boundary is φ. It is formally equal to the product of $\exp(-D(\varphi))$ and $\Delta^{-1/2}$ if Δ is the ζ-function regularization (in the sense of §8 of [G]) of the determinant of the hermitian form

$$-\frac{1}{2\pi^2}\int_R (\nabla f, f)dxdy$$

(For simplicity the π appearing in (4.4) is now absorbed into Δ.) Thus $z(\varphi) = \delta Z(\varphi)$ if $\delta = \Delta^{1/2}\exp(\epsilon)$. Although the value of δ is clearly of interest, I have not tried to calculate it.

If R is the rectangle of sides a and b, the eigenfunctions of this form are $f_{k,l} = \sin\frac{\pi kx}{a}\sin\frac{\pi ly}{b}$ and the eigenvalues (with respect to the second, auxiliary form) are

$$\frac{1}{2\pi^2}(\frac{\pi^2k^2}{a^2} + \frac{\pi^2l^2}{b^2}).$$

Thus

$$\Delta = \prod_{k,l=1}^{\infty}\frac{1}{2\pi^2}(\frac{\pi^2k^2}{a^2} + \frac{\pi^2l^2}{b^2}).$$

The product

$$\prod_{k,l=1}^{\infty}\frac{1}{2a^2} = \big(\frac{1}{2a^2}\big)^{1/4}.$$

(When comparing with [G], observe that $(-\frac{1}{2})^2 = \frac{1}{4}$.) This leaves

$$\prod_{k,l=1}^{\infty}(k + l\frac{ia}{b})(k - l\frac{ia}{b}) = \eta(e^{-a/b})^2$$

if η is the function

$$\eta(q) = q^{1/24}\prod_{m=1}^{\infty}(1-q^m). \tag{4.5}$$

We shall have other regularizations of this type to calculate whose results will be more important to us. I do find it curious that the result is not homogeneous of degree 0 (perhaps because of my inexperience).

5. Dynamics. The dynamics for this model are transparent, and quite analogous to those for percolation. Choose positive integers l and n and divide each side of the square $0 \leq x, y \leq 1$ into intervals of length $1/l$. Take as boundary conditions the set $\Sigma_{l,n}$ of functions $\sigma = \exp(i\varphi/\rho)$ defined by real functions φ that are continuous on the boundary and linear on each of these subintervals and whose values at their endpoints lie in $\{2\pi m/n\rho | m \in \mathbb{Z}\}$. If we demand in addition, and that is certainly reasonable, that the absolute value of the difference between the values of φ at the two endpoints of any interval is bounded (by say $2\pi/\rho$) then $\Sigma_{l,n}$ is a finite set that with increasing l and n approximates the set of continuously differentiable functions on the boundary. The collection Π_l introduced for percolation is replaced by the collection $\mathbf{Z_{l,n}}$ of nonnegative functions on $\Sigma_{l,n}$, two such functions being equivalent if they differ by a factor and the function that is identically 0 being excluded.

The function $z : \sigma \to z(\sigma)$ can be restricted to $\Sigma_{l,n}$ to define an element $z_{l,n}$ in $\mathbf{Z_{l,n}}$. This yields a coherent family from which, I suppose, it is possible to recover the full function z. The transformation $\Theta_{l,n}$ whose fixed point approximates $z_{l,n}$ is clearly at hand.

Observe first of all that there are maps from $\Sigma_{l,2n}$ to $\Sigma_{l,n}$ and from $\Sigma_{2l,n}$ to $\Sigma_{l,2n}$ and thus a map from $\Sigma_{2l,n}$ to $\Sigma_{l,n}$ and hence from $\mathbf{Z_{2l,n}}$ to $\mathbf{Z_{l,n}}$. The second of these two maps is clear because a function φ defining an element of $\Sigma_{l,n}$ also defines an element of $\Sigma_{2l,n}$. Its value at the midpoint of an interval of length $1/l$ is the average of its values at the endpoints. To pass from $\Sigma_{l,2n}$ to $\Sigma_{l,n}$ replace the values of φ at those endpoints where they are equal to $2(2k+1)\pi/2n\rho$ by $2k\pi/n\rho$ or by $2(k+1)\pi/n$. The choice is somewhat arbitrary; the difference between the values at endpoints is not to be increased beyond the allowed size.

To define $\Theta_{l,n}$ we need only define a map from $\mathbf{Z_{l,n}}$ to $\mathbf{Z_{2l,n}}$. Take the square $\square$ of side 1 and divide it into four equal squares $\square_{i,j}$, $1 \le i, j \le 2$, of side $1/2$. Each of these squares may be identified with the original one by the obvious translation and dilation. We divide each side of the small squares into l intervals of length $1/l$. This is compatible with the division of the sides of the large square into intervals of length $1/2l$. If $\sigma \in \Sigma_{2l,n}$ is defined by φ consider collections of functions $\sigma_{i,j} = \exp(\sqrt{-1}\varphi_{i,j}/\rho)$, $1 \le i, j \le 2$ defining boundary conditions on the four smaller squares and such that: 1) $\varphi_{i,j}$ agrees with φ on the intersection of the boundaries of $\square$ and $\square_{i,j}$; 2) $\varphi_{i,j}$ agrees with $\varphi_{i',j'}$ on the intersection of the boundaries of $\square_{i,j}$ and $\square_{i',j'}$. If w lies in $\mathbf{Z_{l,n}}$ map it to the function w' defined by

$$w'(\sigma) = \sum_{\{\sigma_{i,j}\}} \prod_{i,j=1}^{2} w(\sigma_{i,j}), \tag{5.1}$$

the sum running over all the collections just defined. Recall that we are free to divide this sum by a positive factor.

In order to emphasize the analogy with the earlier constructions for percolation, we have taken the values of φ at the endpoints of the intervals to lie in a set that is finite modulo $2\pi\mathbb{Z}/\rho$. Then $\Sigma_{l.n}$ is a finite set. The dynamics are more transparent if we quickly pass to the limit $n \to \infty$ without making any great effort at justification and drop the constraint on the difference between values of φ at consecutive endpoints. Then the collection $\Phi_{l,n}$ of functions φ defining the elements of $\Sigma_{l,n}$ is replaced by the linear space Φ_l of continuous functions on the boundary of the square that are linear on each of the $4l$ subintervals. Taking advantage of the optional factor in the sum of (5.1) we replace it by

$$\int_{\{\varphi_{i,j}\}} \prod_{i,j=1}^{2} w(\sigma_{i,j}). \tag{5.2}$$

Some constraints on w will be necessary to assure the existence of the integral.

To what extent is the restriction of the function z to $\Sigma_l = \{\exp(i\varphi/\rho) | \varphi \in \Phi_l\}$ a fixed point of the transformation (5.2)? The set of collections $\{\varphi_{i,j}\}$ is an affine space, the associated linear space being the set of collections attached to the function on the boundary of $\square$ that is identically 0. Moreover when $w = z$, (5.2) may be written as

$$\int_{\{\varphi_{i,j}\}} \exp(-\sum_{i,j} D(\varphi_{i,j})). \tag{5.3}$$

The form

$$\sum_{i,j} D(\varphi_{i,j}) \tag{5.4}$$

is a positive semidefinite form on the set of all possible collections $\{\varphi_{i,j}\}$. Let $D_l(\varphi)$ be its minimum on the affine space attached to φ. Then apart from a factor that is given by a Gaussian integral, that does not depend on φ, and that may, therefore, be suppressed, the integral (5.3) is simply $\exp(-D_l(\varphi))$.

As a consequence z fails to be a fixed point of (5.2) to the extent that $D_l(\varphi)$ differs from $D(\varphi)$. The difference results from the constraints on the $\varphi_{i,j}$. If we accepted arbitrary continuously differentiable functions for the restrictions of the $\varphi_{i,j}$ to the interior boundaries of $\square_{i,j}$ then the minimum of (5.4) would be achieved by taking $\varphi_{i,j}$ to be the restriction to the boundary of $\square_{i,j}$ of the harmonic extension of φ to the interior of $\square$ and the mimimum would be $D(\varphi)$. Thus we can expect $D_l(\varphi)$ to approach $D(\varphi)$. Whether there is indeed a fixed point of the transformation (5.2) approximately equal to $w_l : \sigma \to \exp(-D_l(\varphi))$ remains to be proved. It is not my purpose here to do this; I only want to stress that for this one, very special, lattice model a dynamical construction analogous to that for percolation is plausible.

6. Conformal field theory. The scaling limits of lattice models are usually defined as for the Ising model in terms of correlation functions. In dimension two these scaling limits are often conformally invariant fields. The close relation between the statistics of the lattice models and the characters (of the Virasoro algebra) attached to the conformally invariant theory is striking and persuasive, especially for the Ising model where all calculations are exact.

The statistics of a lattice model are, for the present purposes, construed as those contained, implicitly or explicitly, in the reduced partition function which it is therefore a matter of representing in terms of the conformal field attached to the model. We could work with a theory on any Riemann surface R (with metric respecting the conformal structure) bounded by a set of *parametrized* simple closed curves $C_1, \ldots, C_r$. It is assumed that the conformal structure on the surface is that appropriate to the given lattice theory. (It is left to the reader's imagination what such a theory might be in the absence of an obvious lattice. There are clearly various possibilities, some more suitable to experimentation than others. For an example in the context of percolation, see [SA].)

The scaling limit is supposed conformally invariant. A conformal theory* is a collection $\{\mathbb{H}_\eta, \mathbb{H}_{\bar\eta}\}$ of pairs of Hilbert spaces on each of whose factors the Virasoro algebra acts. In the notation η is not always explicit. Thus

$$L_n = z^{1-n}\frac{d}{dz} = \frac{e^{-in\theta}}{i}\frac{d}{d\theta}$$

acts on $\mathbb{H}_\eta \otimes \mathbb{H}_{\bar\eta}$ as $L_n \otimes 1$ and $1 \otimes \bar L_n$.

The central charge c is the same for $\mathbb{H}_\eta$ as for $\mathbb{H}_{\bar\eta}$ and independent of η. The highest weights (for all these spaces are highest-weight modules in the sense of [G]) are h_η and $h_{\bar\eta}$. The sesquilinear involution $L \to L^\iota$ that takes L_n to $-L_{-n}$ takes the central element Z to $-Z$ and fixes real vector-fields on the circle. It is supposed (at least by me) that the representation $L \otimes 1 + 1 \otimes \bar L^\iota$ of the Virasoro algebra on $\mathbb{H}_\eta \otimes \mathbb{H}_{\bar\eta}$ extends to an action of diffeomorphisms of the circle.

In a conformal theory a copy $\mathbb{H}^i_\eta$ and $\mathbb{H}^i_{\bar\eta}$ of each $\mathbb{H}_\eta$ and $\mathbb{H}_{\bar\eta}$ is attached to each of the parametrized circles bounding the surface and there are maps, a linear family of them,

$$\otimes_{i=1}^r \mathbb{H}^i_{\eta_i} \to \mathbb{C},$$
$$\otimes_{i=1}^r \mathbb{H}^i_{\bar\eta_i} \to \mathbb{C},$$

but a canonical map γ_R from $\otimes\mathfrak{X}_i$ to $\mathbb{C}$, if

$$\mathfrak{X}_i = \oplus_\eta \mathbb{H}^i_\eta \otimes \mathbb{H}^i_{\bar\eta}.$$

Recall that γ_R is defined for each Riemann surface and that there are a large number of conditions that it must satisfy ([MS]). The spaces $\mathfrak{X}$ are of course all the same, and we denote their standard representative

$$\oplus_\eta \mathbb{H}_\eta \otimes \mathbb{H}_{\bar\eta}$$

by $\mathfrak{X}$.

If $\{\sigma_i | i = 1, \ldots, r\}$ is a collection of boundary conditions (on the curves C_i) for the model then, as for planar domains, we can expect to define the reduced partition function $z(\sigma_1, \ldots, \sigma_r)$. The statistical question is whether there is a map $\sigma \to \xi(\sigma)$ from the pertinent boundary conditions to $\mathfrak{X}$ (or rather to an appropriate completion of $\mathfrak{X}$ in some weak topology) such that for <u>all</u> Riemann surfaces R there is a constant π_R such that for all boundary conditions

$$z(\sigma_1, \ldots, \sigma_r) = \pi_R \gamma_R(\otimes \xi(\sigma_i)). \tag{6.1}$$

Diffeomorphisms of the circle act on boundary conditions and ξ is to be compatible with the two actions, on boundary conditions and on $\mathfrak{X}$. Of course there is, or so it seems to me at present, a good chance that this question has an affirmative response. There are some results in the literature ([C1,Ch]), but for most models, especially –

* I am learning conformal field theory piecemeal. All being well I have not overstepped in these comments the narrow limits of my knowledge in any serious way.

as already observed – for the Ising model, the proper general notion of boundary condition is not immediately manifest, nor is the action of diffeomorphisms on the boundary conditions. The equation (6.1) gives the right side a statistical significance that it does not otherwise have. It cannot be expected that z is conformally invariant but if (6.1) is correct the departure from conformal invariance is all in the constant π_R.

It is an instructive exercise to determine the map ξ for the theory attached to free bosons. The spaces $\mathbb{H}_\eta$ and $\mathbb{H}_{\bar\eta}$ are described in [G]. Recall that the central character $c = 1$. The indices η are parametrized by two integers m and n, both arbitrary, and

$$h_\eta + h_{\bar\eta} = \frac{m^2}{4\rho^2} + n^2\rho^2,$$
$$h_\eta - h_{\bar\eta} = mn.$$

The spaces $\mathbb{H}_\eta$ and $\mathbb{H}_{\bar\eta}$ are the Verma modules with these highest weights. (If h_η or $h_{\bar\eta}$ is $N^2/4$, $N \in \mathbb{Z}$, this assertion may have to be modified, but that is again a fine point and shall be overlooked.) We begin by studying the region R formed not by a planar annulus of center 0, outer radius 1 and inner radius $q < 1$, but by the quotient of the strip between $\Re(z) = \ln q$ and $\Re(z) = 0$ by the translations $z \to z + 2\pi ik$, $k \in \mathbb{Z}$. (Nonetheless I occasionally use a language appropriate to the planar annulus.) This will lead us directly to the only possible definition of ξ for free bosons. Indeed, the calculations below indicate that the existence of ξ satisfying (6.1) for annuli is far from trivial. I take it as real evidence in favour of the views of this essay.

Without having justified it, I take $z(\varphi)$ to be $Z(\varphi)$. (This is presumably the point at which our concrete choice of representative for the conformal class plays a role. Otherwise z and Z may differ by a factor that depends on q. The assumption that they are the same is justified in large part by the success of the calculation. One of my difficulties as I write this essay is that I do not yet understand how regularized determinants behave under conformal transformation. This is presumably related to the distinction between Z and z.) With this assumption,

$$z(\varphi) = \exp(-D(\varphi))\Delta^{-1/2}, \tag{6.2}$$

where Δ is the ζ-function regularization of the determinant for the annulus. A constant independent of q is of no import in this expression for z but the precise dependence on q is critical.

We begin with the calculation of Δ. The eigenfunctions are

$$\sin(\pi lx/\ln q)e^{iky}$$

with eigenvalues $(k^2 + l^2\frac{\pi^2}{\ln^2 q})/2\pi$. Thus if $\tau' = -\pi i/\ln q$ (so that $e^{\pi i\tau} = q$ if $\tau = -1/\tau'$) then

$$\Delta = \prod_{k=-\infty}^{\infty}\prod_{l=1}^{\infty}(k + l\tau')(k - l\tau').$$

(The factor 2π in the denominator does not contribute to the regularized determinant.) Using the results of [G,§8] in which the square of the right hand side is calculated, I obtain

$$\Delta = \eta^2((q')^2), \qquad q' = e^{\pi i\tau'}.$$

The functional equation of η is however

$$\sqrt{-i\tau}\,\eta(q^2) = \eta((q')^2),$$

so that

$$\Delta^{-1/2} = (-i\tau)^{-1/2}\eta^{-1}(q^2). \tag{6.3}$$

It is, of course, not φ that is of ultimate concern, but $\sigma = \exp(i\varphi/\rho)$. Thus we can add integral multiples of $2\pi\rho$ to φ without affecting σ. The function φ is in reality two functions, φ_1 on the inner boundary of the annulus and φ_2 on the outer boundary. Neither has to be periodic, but

$$\varphi_i(\theta + 2\pi) = \varphi_i(\theta) - 2\pi n\rho,$$

the integer n being the same for both the inner and outer boundaries, for otherwise it would be impossible to extend φ into the interior of the annulus, and $Z(\varphi)$ would be 0, the corresponding sum (or integral) being empty. Moreover we may add different integral multiples of $2\pi\rho$ to φ_1 and to φ_2, say $2m_1\pi\rho$ and $2m_2\pi\rho$. One is irrelevant because we can remove it by adding the same multiple of $2\pi\rho$ to the extension f of φ into the interior of the annulus, and this has no effect on $s = \exp(if/\rho)$. (More generally, adding a common constant to the functions φ_1 and φ_2 has no effect on the partition function.) The difference $m = m_2 - m_1$ must, however, be taken into account, so that

$$z(\sigma) = \sum_{m,n} Z(\varphi_1 + a\ln z + b\ln\bar{z}, \varphi_2 + a\ln z + b\ln\bar{z}).$$

Here φ is defined by

$$\begin{aligned}\varphi_1(\theta) &= \sum_{k\neq 0} a_k e^{ik\theta}, \qquad a_{-k} = \bar{a}_k,\\ \varphi_2(\theta) &= \sum_{k\neq 0} b_k e^{ik\theta}, \qquad b_{-k} = \bar{b}_k.\end{aligned}$$

and

$$\begin{aligned}-a\ln q - b\ln q &= x + 2\pi m\rho,\\ (a-b) &= in\rho.\end{aligned}$$

The number x is determined, even though only modulo $2\pi\rho\mathbb{Z}$, by the boundary conditions σ. These equations imply that $b = \bar{a}$.

We extend φ_1 to the annulus as a harmonic function $\tilde{\varphi}_1$ that is 0 on the outer boundary, and φ_2 to the annulus as a harmonic function $\tilde{\varphi}_2$ that is 0 on the inner boundary. The function $a\ln z + b\ln\bar{z}$ is of course already harmonic in the whole annulus and is real. The function $\tilde{\varphi}$ will then be the sum of these three functions.

There are simple formulas for $\tilde{\varphi}_1$ and $\tilde{\varphi}_2$.

$$\begin{aligned}\tilde{\varphi}_1(z) &= \sum_{k\neq 0} \frac{a_k}{q^k - \frac{1}{q^k}}\{z^k - \bar{z}^{-k}\},\\ \tilde{\varphi}_2(z) &= \sum_{k\neq 0} \frac{b_k}{\frac{1}{q^k} - q^k}\{\big(\frac{z}{q}\big)^k - \big(\frac{\bar{z}}{q}\big)^{-k}\}\end{aligned}$$

The appropriate form is

$$D(\varphi) = \frac{2}{\pi}\int \frac{\partial\tilde{\varphi}}{\partial z}\frac{\partial\tilde{\varphi}}{\partial\bar{z}}dxdy.$$

Since

$$\begin{aligned}\frac{2}{\pi}\int d\theta\int_q^1 r^{2k}\frac{dr}{r}dr &= \frac{2(1-q^{2k})}{k},\\ \frac{2}{\pi}\int d\theta\int_q^1 \frac{dr}{r} &= 4\ln 1/q,\end{aligned}$$

we obtain $z(\sigma)$ as the product of two terms that are treated quite differently. The first is a term that is independent of m and n,

$$\prod_{k=1}^{\infty}\exp(-2k(a_k a_{-k}\frac{1+q^{2k}}{1-q^{2k}} - a_k b_{-k}\frac{2q^k}{1-q^{2k}} - b_k a_{-k}\frac{2q^k}{1-q^{2k}} + b_k b_{-k}\frac{1+q^{2k}}{1-q^{2k}})). \tag{6.4}$$

(The signs here are important; so I give a couple of the calculations. First of all, for $b_k b_{-k}$ we obtain from $\frac{\partial\tilde{\varphi}_1}{\partial z}\frac{\partial\tilde{\varphi}_1}{\partial\bar{z}}$

$$-\frac{2}{\pi}k^2\frac{1}{\frac{1}{q^k}-q^k}\frac{1}{\frac{1}{q^{-k}}-q^{-k}}\frac{1}{q^k}\frac{1}{q^k}\int_0^{2\pi}d\theta\int_q^1 r^{2k-1}dr = 2k\frac{1}{1-q^{2k}}$$

and, by symmetry,

$$-2k\frac{1}{1-q^{-2k}} = 2k\frac{q^{2k}}{1-q^{2k}}.$$

The two expressions are to be added together. For $b_k a_{-k}$ we obtain from $\frac{\partial\tilde\varphi_1}{\partial z}\frac{\partial\tilde\varphi_2}{\partial\bar z}$

$$-\frac{2}{\pi}k^2\frac{1}{\frac{1}{q^k}-q^k}\frac{1}{q^{-k}-\frac{1}{q^{-k}}}\frac{1}{q^k}\int_0^{2\pi}d\theta\int_q^1 r^{2k-1}dr = -2k\frac{q^k}{1-q^{2k}}.$$

From $\frac{\partial\tilde\varphi_1}{\partial\bar z}\frac{\partial\tilde\varphi_2}{\partial z}$ we obtain

$$-2k^2\frac{1}{\frac{1}{q^k}-q^k}\frac{1}{q^{-k}-\frac{1}{q^{-k}}}\frac{1}{q^{-k}}(1-q^{-2k}) = -2k\frac{q^k}{1-q^{2k}}.$$

Once again these two expressions are to be added together.)

Observing that

$$a = \frac{1}{2}\{\frac{x+2\pi m\rho}{\ln(1/q)} - in\rho\},\qquad b = \frac{1}{2}\{\frac{x+2\pi m\rho}{\ln(1/q)} + in\rho\}$$

so that

$$ab = \frac{1}{4}\{\frac{(x+2\pi m)^2}{\ln(1/q)^2} + n^2\rho^2\},$$

we obtain for the second term

$$\sum_{m,n=-\infty}^{\infty}\exp(-\ln(1/q)\{\frac{(x+2\pi m\rho)^2}{\ln^2(1/q)} + n^2\rho^2\}).$$

This is itself a product

$$\{\sum_{m=-\infty}^{\infty}\exp(-\frac{(x+2\pi m\rho)^2}{\ln(1/q)}\}\{\sum_{n=-\infty}^{\infty}\exp(-n^2\rho^2\ln(1/q)\}.$$

The second factor is $\sum q^{n^2\rho^2}$. To the first we apply the Poisson summation formula to obtain

$$\frac{\sqrt{|\tau|}}{2\rho}\sum_{m=-\infty}^{\infty}e^{imx/\rho}q^{\frac{m^2}{4\rho^2}}.$$

The product is therefore

$$\frac{\sqrt{|\tau|}}{2\rho}\sum_{m,n}e^{imx/\rho}q^{\frac{m^2}{4\rho^2}+n^2\rho^2}. \tag{6.5}$$

I observe first of all that the factor $1/\sqrt{-i\tau}$ of (6.3) multiplied with the factor $\sqrt{|\tau|}$ of (6.5) is 1. Thus there is a constant factor $1/2\rho$ left over. It is of no importance.

Each of the spaces $\mathbb{H}_\eta$ and $\mathbb{H}_{\bar\eta}$, and therefore the space $\mathfrak{X}$ itself, is equipped with an invariant bilinear form, the Shapovalov form. If I am not mistaken, the map γ of conformal field theory is for an annulus of invariant q given by

$$q^{-c/12}\sum_\eta(q^{L_0}\otimes q^{\bar L_0}\mathfrak{x}_1^\eta,\mathfrak{x}_2^\eta),$$

if $\mathfrak{x}_1$ lies in the space $\mathfrak{X}_1$ attached to the inner boundary and $\mathfrak{x}_2$ lies in the space $\mathfrak{X}_2$ attached to the outer boundary. (Observe that we are implicitly using the natural parametrizations of the two boundaries by θ.) Recall that for free bosons $c=1$. Moreover, by (4.5), the factor $q^{-1/12}$ is present in the factor $\eta^{-1}(q^2)$ appearing in (6.3).

Our problem is therefore (we suppress the constant for it can be incorporated into the formulas at any time) to show that

$$\sum_{m,n} e^{imx/\rho} q^{\frac{m^2}{4\rho^2}+n^2\rho^2} \prod_{k=1}^{\infty} \frac{\exp(-2k(a_k a_{-k}\frac{1+q^{2k}}{1-q^{2k}} - a_k b_{-k}\frac{2q^k}{1-q^{2k}} - b_k a_{-k}\frac{2q^k}{1-q^{2k}} + b_k b_{-k}\frac{1+q^{2k}}{1-q^{2k}}))}{1-q^{2k}}, \tag{6.6}$$

in which we have combined the infinite products of (4.5) and (6.4), can be written as

$$\sum_{\eta} (q^{L_0} \otimes q^{\bar{L}_0} \mathfrak{x}^{\eta}(\varphi_1), \mathfrak{x}^{\eta}(\varphi_2)). \tag{6.7}$$

The function $\mathfrak{x}^{\eta}(\psi)$ is now understood to be defined for functions on the circle. It is applied to φ_1 and φ_2.

We observe first of all ([G]) that the eigenvalue structure of $q^{L_0} \otimes q^{\bar{L}_0}$ and of $e^{i\theta L_0} \otimes e^{-i\theta \bar{L}_0}$ on $\mathbb{H}_\eta \otimes \mathbb{H}_{\bar{\eta}}$, $\eta = \eta(m,n)$ is quite simple. There are first of all common factors $q^{\frac{m^2}{4\rho^2}+n^2\rho^2}$, which appears clearly in (6.6), and $e^{imn\theta}$, which also appears, as we shall see, but not so clearly. Secondly, because each of the modules $\mathbb{H}_\eta$ and $\mathbb{H}_{\bar{\eta}}$ is a Verma module, what is left is the tensor product from $k=1$ to infinity of modules, each of which is itself a tensor product of two modules, one on which $qe^{i\theta}$ acts with weights $q^{lk}e^{ilk\theta}$, $l = 0, \ldots, \infty$, of multiplicity one, and one on which $qe^{i\theta}$ acts with weights $q^{lk}e^{-ilk\theta}$, $l = 0, \ldots, \infty$, again of multiplicity one.

In general if we have a space on which $\mathbb{C}^\times \times \mathbb{C}^\times$ (conceived here as containing $qe^{i\theta} \times qe^{-i\theta}$) acts and two non-degenerate forms $(*,*)_1$ and $(*,*)_2$ such that the eigenspaces of $\mathbb{C}^\times \times \mathbb{C}^\times$ are mutually orthogonal with respect to both forms then there is a transformation J commuting with the action of this group such that

$$(x,y)_1 = (Jx, Jy)_2$$

Therefore

$$((z_1 \times z_2)x, y)_1 = (J(z_1 \times z_2)x, Jy)_2 = ((z_1 \times z_2)Jx, Jy)_2.$$

Consequently it suffices to define the function $\mathfrak{x}^{\eta}(\psi)$ so that (6.7) is satisfied for some more convenient form on $H_\eta \otimes H_{\bar{\eta}}$ than the Shapovalov form. We use the tensor product structure, respected by the action, dropping the common factor $q^{h_\eta}\bar{q}^{h_{\bar{\eta}}}$. (I have always taken q to be real and shall continue by and large to do so, but it is often convenient to write q for $qe^{i\theta}$ and $\bar{q}$ for $qe^{-i\theta}$.) Each factor of the tensor product can be regarded as a space of infinite matrices X on which $q \times \bar{q}$ acts as

$$X \to \begin{pmatrix} 1 & 0 & 0 & \ldots \\ 0 & Q & 0 & \ldots \\ 0 & 0 & Q^2 & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} X \begin{pmatrix} 1 & 0 & 0 & \ldots \\ 0 & \bar{Q} & 0 & \ldots \\ 0 & 0 & \bar{Q}^2 & \ldots \\ \vdots & \vdots & \vdots & \ddots, \end{pmatrix}$$

with $Q = q^k$. The appropriate form is

$$\mathrm{Trace}(XY).$$

I have already observed that (6.6) comes as a sum over the m and n parametrizing the η. I also observe that there is a redundancy in the parametrization that is at first curious, for the pairs (m,n) and $(-m,-n)$ yield the same η and $\bar{\eta}$. Our construction of ξ does not assure that ξ commutes with the action of all diffeomorphisms, and I have not yet attempted to prove that it does. This would be beyond my present understanding of the theory. The construction is, however, guided by the condition that it be compatible with independent rotations of the inner and outer boundaries. Rotating the inner boundary by θ_1 and the outer by θ_2 introduces (among other things) a factor $e^{imn(\theta_1-\theta_2)}$ (the sign comes from a contragredient action) into the term of (6.7) attached to m and n. On the other hand these rotations applied to $\varphi_i + a\ln z + b\ln\bar{z}$ replace x by $x + (a-b)(\theta_2-\theta_1)i = x + (\theta_1-\theta_2)n\rho$. This introduces the same factor $e^{im(\theta_1-\theta_2)n}$ into (6.6), one of the first signs that we are on the right track. The number x is the difference, modulo $2\pi\rho\mathbb{Z}$, between the averages x_1 and x_2 of φ on the outer boundary and the inner boundary. Thus the factor $e^{imx/\rho}$ is the product of $e^{imx_1/\rho}$ and $e^{-imx_2/\rho}$, one to be incorporated into $\mathfrak{x}^{\eta}(\varphi_1)$

and the other into $\mathfrak{r}^\eta(\varphi_2)$. This process is asymmetrical and is only possible because of the redundancy already observed. Replacing (m,n) by $(-m,-n)$ reverses the roles of φ_1 and φ_2.

This leaves for each k and $Q = q^k$ the expression

$$\frac{e^{-2k(a_+a_-(1+Q^2)-2a_+b_-Q-2b_+a_-Q+b_+b_-(1+Q^2))/(1-Q^2))}}{1-Q^2},$$

in which I have written a_+ and a_- for a_k and a_{-k} and b_+ and b_- for b_k and b_{-k} There is an obvious factor of this expression,

$$e^{-2ka_+a_-}e^{-2kb_+b_-}, \tag{6.8}$$

that is itself a product. Thus only

$$\frac{e^{-4k(a_+a_-Q^2-a_+b_-Q-b_+a_-Q+b_+b_-Q^2)/(1-Q^2)}}{1-Q^2} \tag{6.9}$$

matters. This is to be written as the trace of

$$\begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & Q & 0 & \dots \\ 0 & 0 & Q^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} A \begin{pmatrix} 1 & 0 & 0 & \dots \\ 0 & Q & 0 & \dots \\ 0 & 0 & Q^2 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} B,$$

where A is a function of a_+ and a_- and B is the same function of b_+ and b_-. The factor $-4k$ in the exponent is a nuisance as are the signs within the parentheses, but the factor and the signs can be removed by incorporating a factor $2i\sqrt{k}$ into each of a_+, a_- and a factor $-2i\sqrt{k}$ into b_+ and b_- (an asymmetric procedure), thereby replacing (6.9) by

$$\frac{e^{(a_+a_-Q^2+a_+b_-Q+b_+a_-Q+b_+b_-Q^2)/(1-Q^2)}}{1-Q^2} \tag{6.10}$$

The numerator of (6.10) is best expressed as a product of two factors. We expand them, obtaining

$$\sum\nolimits_{i=0}^\infty \frac{(a_+a_-+b_+b_-)^i}{i!}\frac{Q^{2i}}{(1-Q^2)^i},$$
$$\sum\nolimits_{j,k=0}^\infty \frac{(a_+b_-)^j(b_+a_-)^k}{j!k!}\frac{Q^{j+k}}{(1-Q^2)^{j+k}}.$$

Then we group together the factors $(1-Q^2)$ in the denominators, incorporate the denominator of the original expression, and expand again

$$\frac{1}{(1-Q^2)^{i+j+k+1}} = \sum_{l=0}^\infty \frac{(i+j+k+l)!}{(i+j+k)!l!}Q^{2l}.$$

Thus (6.10) is equal to

$$\sum_{i,j,k,l=0}^\infty \frac{(a_+a_-+b_+b_-)^i}{i!}\frac{(a_+b_-)^j(b_+a_-)^k}{j!k!}(i+j+k+l)!(i+j+k)!l!Q^{2i+j+k+2l} \tag{6.11}$$

This is to be equal to

$$\sum_{m,n=0}^{\infty} A_{m,n} Q^{m+n} B_{n,m}, \tag{6.12}$$

where $A_{m,n}$ is a function of a_+ and a_- and $B_{m,n}$ is (almost) the same function of b_+ and b_-. (The indices m and n are no longer the parameters of the η describing the sectors of the theory!) Because the map ξ is supposed to commute with the action of the group of diffeomorphisms of the circle, and therefore in particular with rotations, the function $A_{m,n}$, a polynomial in a_+ and a_-, is a finite linear combination of the monomials $a_+^p a_-^q$, $p-q=m-n$.

Thus $A_{m,n}B_{n,m}$ is the sum of all terms in (6.11) for which $j-k=m-n$ and $2i+j+k+2l=m+n$. By symmetry it suffices to consider the case $m \geq n$. Fixing m and n we extract from each term in the sum the factor $(a_+b_-)^{m-n}$, an expression that is obviously a function of (a_+, a_-) times a function of (b_+, b_-). What remains we write as a function of $x = a_+a_-$ and $y = b_+b_-$,

$$S_{m,n}(x,y) = \sum_{i+k+l=n} \frac{(x+y)^i (xy)^k}{i!(m-n+k)!k!} \frac{(m+k)!}{(m-n+i+2k)!l!}.$$

All indices in the summation are supposed positive or zero. This is a symmetric function of x and y. Set

$$R_{m,n}(x) = S_{m,n}(x,0) = \sum_{i+l=n} \frac{x^i}{i!(m-n)!} \frac{m!}{(m-n+i)!l!}.$$

and

$$T_{m,n} = R_{m,n}(0) = S_{m,n}(0,0) = \frac{m!}{(m-n)!^2 n!}.$$

Then the necessary and sufficient condition for $A_{m,n}$ to exist is that

$$S_{m,n}(x,y)T_{m,n} = R_{m,n}(x)R_{m,n}(y), \tag{6.13}$$

and then

$$\begin{aligned} A_{m,n}(a_+,a_-) &= a_+^{m-n} R_{m,n}(x)/\sqrt{T_{m,n}}, \quad m \geq n \\ A_{m,n}(a_+,a_-) &= (-1)^{m-n} a_-^{n-m} R_{n,m}(x)/\sqrt{T_{n,m}}, \quad n \geq m. \end{aligned}$$

The sign results from the asymmetric treatment of (a_+, a_-) and (b_+, b_-).

To see what the identity amounts to we compare coefficients of monomials in x and y. After a common factor of $(m-n)!^2$ is removed from the denominators, (6.13) reduces to the identity of

$$\frac{m!}{n!} \sum_k \frac{1}{(j_1-k)!(j_2-k)!(m-n+k)!k!} \frac{(m+k)!}{(m-n+j_1+j_2)!(n+k-j_1-j_2)!},$$

in which the variable of summation k runs from $\max\{0, j_1+j_2-n\}$ to $\min\{j_1, j_2\}$, and

$$\frac{m!}{j_1!(m-n+j_1)!(n-j_1)!} \frac{m!}{j_2!(m-n+j_2)!(n-j_2)!},$$

for each fixed j_1 and j_2 less than or equal to n.

If $j_1 = 0$ then the sum over k contains a single term, with $k=0$; if $j_1 = n$ it contains a single term, with $k = j_2$. In both cases the identity is easily verified. If $j_1 = 1$ and $j_2 < n$ the sum runs from 0 to 1 and the identity is again easily verified. In general, as Fan Chung immediately recognized, it is the identity of Saalschütz. (Substitute $r=m$, $s=n$, $n=j_2$, $m=j_1+m-n$ in formula (5.28) of [GKP].)

I give some simple examples. The degree of $R_{m,n}$ is n, or more generally, $\min\{m,n\}$.

$$\begin{aligned} R_{0,0}(x) &= 1, R_{1,1}(x) &= x+1, R_{2,2}(x) &= \frac{1}{2}x^2+2x+1, \\ R_{0,2}(x) &= \frac{1}{2}, R_{1,3}(x) &= \frac{1}{2}x+\frac{3}{2}, R_{2,4}(x) &= \frac{1}{4}x^2+2x+3. \end{aligned}$$

An example of higher order is

$$R_{5,9}(x) = \frac{1}{2880}x^5 + \frac{1}{64}x^4 + \frac{1}{4}x^3 + \frac{7}{4}x^2 + \frac{21}{4}x + \frac{21}{4}.$$

Another is $R_{8,15}(x)$ which is

$$\frac{1}{203212800}x^8 + \frac{1}{1693440}x^7 + \frac{1}{34560}x^6 + \frac{13}{17280}x^5 + \frac{13}{1152}x^4 + \frac{143}{1440}x^3 + \frac{143}{288}x^2 + \frac{143}{112}x + \frac{143}{112}.$$

Once ξ is defined, the question arises whether (6.1) is valid for other Riemann surfaces than the annulus. At the moment, I am in no position to answer it, except for the disk, and even for the disk only partially. So far as I can tell the only intelligent definition of γ for the disk is to take the product with the highest-weight vector in the vacuum sector – the space $\mathbb{H}_\eta \otimes \mathbb{H}_{\bar\eta}$ with $h_\eta = h_{\bar\eta} = 0$ – which is always present. For free bosons this corresponds to the parameters $m = n = 0$.

Take the disk to be of radius 1 with center at the origin. The various sectors other than the vacuum are no longer pertinent. Except for an inessential additive constant the function φ is determined by σ and the factor $Z(\varphi)$ is easily calculated.

Let

$$\varphi(\theta) = \sum_{k=-\infty}^{\infty} a_k e^{ik\theta}.$$

The harmonic extension $\tilde\varphi$ of φ to the disk is

$$a_0 + \sum_{k=1}^{\infty} a_k z^k + \sum_{k=1}^{\infty} a_{-k}\bar z^k.$$

The constant a_0 has no effect on the partition function and

$$D(\varphi) = \sum_{k=1}^{\infty} 4k^2 a_k a_{-k} \int_0^1 r^{2k-1}dr = \sum_{k=1}^{\infty} 2k a_k a_{-k}.$$

Thus

$$\exp(-D(\varphi)) = \prod_{k=1}^{\infty} e^{-2ka_k a_{-k}},$$

and this is exactly the contribution of the factors (6.8) to $\xi(\varphi)$. We also have the product over all k of the coefficients $A_{0,0}$, but they are all 1. This leaves a constant. Moreover, the square root of the regularized determinant for the Laplacian has to be taken into account, as does the discrepancy, whatever it is, between $z(\varphi)$ and $Z(\varphi)$. These are numbers that are independent of φ but I have not yet calculated them.

7. Cylinders. These calculations serve, in my view, not only to give some credibility to the statistical question but also to clarify the dynamical problem. The dynamical problem was formulated in terms of a space Σ, or rather of approximating spaces Σ_l, all of them introduced in terms of the model. Since, however, we are looking for something universal, they should have a universal significance. A natural hope would be that the map ξ took the boundary conditions for any particular model into a subset of some universal set $\mathfrak{S} \subset \mathfrak{X}$, and that it is this subspace on which the *true* Θ acts. One might hope at the same time that the coarsenings (or ultraviolet cutoffs) necessary for the definition of the finite approximations had some meaning in $\mathfrak{S}$. This is certainly farfetched, but is suggested by the calculations for free bosons.

A natural "multiplicative" coarsening in a vector space X of infinite dimension is a decomposition $X = X_1 \otimes X_2$, with X_1 finite-dimensional, together with the implicit "projection" onto X_1, or, better, a choice of one-dimensional subspace of X_2. A natural "linear" coarsening for functions expanded into Fourier series is to project on the terms that oscillate slowly. (This is of course approximately the same as projecting onto the earlier space of piecewise linear functions.) For free boson fields the linear coarsening of φ and the associated multiplicative coarsening of σ corresponds to the first coarsening on $\xi(\sigma)$ – defined by the tensor product over k, in which there is one factor for each frequency.

The dynamical question was formulated in terms of four squares, because that is the simplest form in which renormalization appears. It is, however, to be expected that the function z has many other transformation properties. For example, rather than pasting together four squares, we could paste together two annuli of invariants q_{12} and q_{23} to obtain one of invariant $q_{12} = q_{12}q_{23}$. Then (5.2) would be replaced by a pasting along the boundary. The boundary condition σ_{12} would be made up of two pieces, σ_1 and σ_2 at the two ends, and σ_{23} of two pieces σ_2 and σ_3

We would expect to be able to combine

$$z(\sigma_{12}) = q_{12}^{-c/12} \sum_\eta (q_{12}^{L_0} \otimes q_{12}^{\bar{L}_0} \mathfrak{x}^\eta(\sigma_1), \mathfrak{x}^\eta(\sigma_2)) \tag{7.1}$$

and

$$z(\sigma_{23}) = q_{23}^{-c/12} \sum_\eta (q_{23}^{L_0} \otimes q_{23}^{\bar{L}_0} \mathfrak{x}^\eta(\sigma_2), \mathfrak{x}^\eta(\sigma_3)) \tag{7.2}$$

into

$$z(\sigma_{13}) = q_{13}^{-c/12} \sum_\eta (q_{13}^{L_0} \otimes q_{13}^{\bar{L}_0} \mathfrak{x}^\eta(\sigma_1), \mathfrak{x}^\eta(\sigma_3)) \tag{7.3}$$

by some kind of gluing along the frontier. The factors before the first two sums certainly combine to give that before the third.

Consider the factors $e^{im_{12}x_{12}}/\rho$ and $e^{im_{23}x_{23}}/\rho$. In the pasting the only condition on x_{12} and x_{23}, both defined in any case only modulo $2\pi\rho\mathbb{Z}$, is that their sum be x_{13}. Thus $x_{23} = x_{13} - x_{12}$. The first step in the process analogous to (5.2) would be to integrate over x_{12} in the product of (7.1) and(7.2). The result is to elimate all terms of the expanded product in which the two indices m are not the same. The same n is associated to all φ_i. In other words the $true$ spaces Σ naturally break up into a disjoint union of subsets indexed by $n \in \mathbb{Z}$ and compatible boundary conditions have to lie in the same subset.

This leaves, for each n, a pasting of individual terms in the sum of (6.6) for $q = q_{12}$ with the corresponding term for $q = q_{23}$ to obtain that for $q = q_{13}$. Certainly $x_{13} = x_{12} + x_{23}$ so that the first factors are compatible. So are the second. This leaves the terms of the products. We fix a k, and treat them individually. Then the first relevant parameters are the a_k and a_{-k} for the first annulus, and we set $a_+ = 2\sqrt{k}a_k$ and $a_- = 2\sqrt{k}a_{-k}$. This is not the same assignment as before. We now have $a_- = \bar{a}_+$,and there will be a lot of signs in the formulas. The b_k and b_{-k} for the first annulus are the same as the a_k and a_{-k} for the second annulus, and we set $z = 2\sqrt{k}b_k$, $\bar{z} = 2\sqrt{k}b_{-k}$. Finally we replace the b_k and b_{-k} for the second annulus by $b_+ = 2\sqrt{k}b_k$, $b_- = 2\sqrt{k}b_{-k}$. The temporary ambiguity in notation has been removed. I observe also that a_- and b_- are redundant.

The contributions from the k-th factor are, first of all,

$$e^{-a_+a_-/2}e^{-b_+b_-/2}e^{-z\bar{z}}$$

which when the changes in notation are taken into account is seen to be the factor (6.8) for the large annulus times $e^{-z\bar{z}}$. (I am ignoring the constants that appear in $\mathfrak{x}^\eta$.)

To explain the next step the representation of the k-th factor of the tensor product by matrices is inconvenient. We think rather in terms of a space with basis $u_1, u_2, \ldots$, a function

$$u(z) = \sum_i \alpha_i(z) u_i$$

with values in it and two operators Q_{21} and Q_{32}, and consider

$$\int_z (Q_{21}u(a_+), u(z))(Q_{32}u(z), u(b_+)),$$

which we want to equal

$$(Q_{32}Q_{21}u(a_+), u(b_+)).$$

Suppose that for each i there is a unique i' such that $(u_i, u_j) = \delta_{i',j}$. Then what we need is that

$$\int_z \alpha_i(z)\alpha_j(z) = \delta_{i',j}$$

(perhaps up to a common constant).

In the present context i is a pair (m, n), locating the entry of a matrix, and i' is (n, m). The integral is (for $j = (n', m')$)

$$\int e^{-z\bar{z}} A_{m,n}(z, \bar{z}) A_{n',m'}(z, \bar{z}) dz d\bar{z}.$$

It is clear that this integral is 0 unless $m - n = m' - n'$. When this equality is satisfied it becomes for $m \geq n$ (it clearly suffices to treat this case)

$$2\pi \int_0^\infty e^{-x} R_{m,n}(-x) R_{m',n'}(-x) x^{m-n+1} \frac{dx}{x} / \sqrt{T_{m,n}T_{m',n'}}.$$

Dropping the factor 2π, which is of little interest, we compute this integral explicitly as a sum

$$\sum_{j=0}^{n} \sum_{j'=0}^{n'} (-1)^{j+j'} \frac{m!}{j!(m-n+j)!(n-j)!} \frac{m'!}{j'!(m'-n'+j')!(n'-j')!} (m-n+j+j')! \frac{\sqrt{n!n'!}}{\sqrt{m!m'!}}$$

This expression is equal to $\delta_{m,m'}\delta_{n,n'}$. To verify this, suppose without loss of generality that $n \leq n'$ and recall that $m - n = m' - n'$ so that $\delta_{m,m'}\delta_{n,n'} = \delta_{n,n'}$. Apply the formula (5.24) of [GKP] with $s = n = m' - n' + j$, $l = m'$, and $m = m - n$ to the sum over j'. Since $\mathfrak{n} - \mathfrak{l} = j - n'$ the result is 0 unless $j = n'$ and this is only possible for $n = n'$. If $n = n'$ so that $m = m'$ the sum over j reduces to a single term that is equal to 1. The functions $A_{m,n}(iz, i\bar{z})$ thus form a collection of Hermite polynomials in the plane.

8. Statistics. The attempt to find an intuitively perusasive extension of the notions of this essay to the Ising model is confronted with many obstacles, to some of which I have already alluded. There are others. The approximate dynamics, formulated in terms of a square so that the notion of fixed points was meaningful, is also implicit in the construction of the partition function for the composite of two cylinders from the partition function for the two constitutent cylinders.

For the Ising model all reflections have to begin with the lattice, and the notion of boundary condition that at first presents itself is that defined by a magnetic field on the boundary whose strength diminishes with the mesh. On the other hand in the context of cylindical lattices the gluing, in principle a direct matter, has to be a sum over typical distributions of spins along the bounding edge. If the cylinder is given by sites at (x, y), $1 \leq x \leq M$,

$0 \leq y \leq N$ with the rows at height 0 and N identified then one obvious parameter is the scale of the distribution of the random variable

$$S_x = \sum_{y=1}^{N} s(x,y) \tag{8.1}$$

either for $x = 1, M$ or for $1 << x << M$. The behaviour at the two boundaries is quite different than that for an interior column; and both are pertinent for the gluing. They need to be reconciled. The results of [MW] suggest that the shortest scale on the boundary is $N^{1/2}\ln^{1/2} N$ but that the shortest scale in the interior is $N^{7/8}$. There is, as is to be expected, more conformity where the particles are completely surrounded by their fellow-creatures than on the fringes.

As it turns out the statistics for free bosons behave in the same way. For the Ising model the value of $s(x,y)$ is always ± 1 but for free bosons $s(x,y) = \exp(if(x,y)/\rho)$ can be any number of absolute value 1. The random variable S_x may still be introduced by (8.1), although its physical significance is not patent; and simple informal calculations show that it behaves as for the Ising model. The difficulty, at present insurmountable, is that the measure space approximated either by functions on a column of the discrete cylindrical lattice or, as in the previous section, by low wave-number approximations to functions on the circle is defined not by the functions σ but by the functions φ. For the latter the statistics behave in a coherent manner through the approximation; for the former they do not; and the analogue of φ for the Ising model is not immediately at hand.

Let $\beta = 1/\rho^2$. Then for free bosons the scale at the boundary is:

B.1 $N^{1-\beta/2}$ if $\beta < 1$;
B.2 $N^{1/2}\ln^{1/2} N$ if $\beta = 1$;
B.3 $N^{1/2}$ if $\beta > 1$.

In the interior it is:

I.1 $N^{1-\beta'/2}$ if $\beta' = \beta/2 < 1$;
I.2 $N^{1/2}\ln^{1/2} N$ if $\beta' = 1$;
I.3 $N^{1/2}$ if $\beta' > 1$.

If ρ is so chosen that $\beta = 1$ then this behaviour is similar to that of the Ising model except that $7/8$ is replaced by $3/4$.

The calculations are worth rehearsing. They can be made for the lattices, where they are in principle rigorous, although my treatment will be cavalier; or they can be made informally after the passage to a scaling limit. I begin with the lattice and in the interior.

Since the column is deep in the interior and thus presumably affected only weakly by the edges, I prefer for the sake of the calculation to replace the cylinder by a torus and to work with boundary conditions periodic in both directions, adding a column at $x = 0$ to be identified with that at $x = M$. Consider $S = S_0$. It is a complex random variable and its distribution is manifestly invariant under rotation because $H(f)$ does not change if a constant is added to f. The scale of S is taken to be the square root of the expectation of $|S|^2$, thus to be the square root of

$$\sum_{m,n=1}^{N} E(s(0,m)\bar{s}(0,n)) = N\sum_{n=1}^{N} E(s(0,0)\bar{s}(0,n)). \tag{8.2}$$

The terms of this sum are

$$E(\exp(i(f(0,0) - f(0,n))/\rho) = \frac{\int(\exp(i(f(0,0) - f(0,n))/\rho - H(f))}{\int \exp(-H(f))}, \tag{8.3}$$

in which $H(f)$ is again defined by (4.3), nearest neighbours being taken on the torus.

The denominator being a Gaussian integral, the quotient is a typical characteristic function and easily evaluated. A little care has to be taken, however, with the definition because H is degenerate. It is useful not to work with the usual coordinates for the functions f, thus its values, but with the values of its (two-dimensional) Fourier transform,

$$\tilde{f}(k) = 1/\sqrt{MN}\sum e^{2\pi i k\cdot p} f(p).$$

Of course $\tilde{f}(-k)$ is the complex conjugate of $f(k)$ and the Lebesgue measure defining the Gaussian integral is associated to $1/2\sum \tilde{f}(k)\tilde{f}(-k)$. The form itself becomes

$$H(\tilde{f}) = \frac{2}{\pi}\pi\sum_{u,v}(\sin^2(\pi u) + \sin^2(\pi v))\tilde{f}(u,v), \tag{8.4}$$

with $u = m/M$, $0 \le m < M$, $v = n/N$, $0 \le n < N$. The integral is taken over the space of functions $\tilde{f}$ that vanish at 0. On this space H is invertible with inverse G.

The Fourier transform $\epsilon_{m,n}$ of the function $(\delta_{(0,m)} - \delta_{(0,n)})/\rho$ concentrated at the two points $(0,m)$ and $(0,n)$ is

$$(e^{2\pi i m v} - e^{2\pi i n v})/\rho\sqrt{MN},$$

and (8.3) is $\exp(-\frac{1}{4}G(\epsilon_{0,n}))$. Clearly,

$$\exp(-\frac{1}{4}G(\epsilon_{0,n})) = \exp(-\frac{\pi}{8\rho^2 MN}\sum_{k=(u,v)\neq 0}|1-e^{2\pi i n v}|^2/(\sin^2(\pi u)+\sin^2(\pi v))).$$

(The factor $8 = 2\cdot 4$, the 2 coming from (8.4) and the 4 from the usual formula for characteristic functions with respect to Gaussian integrals.) Since M is sent to ∞ before N the sum can without any disquiet be replaced by an integral to obtain

$$\exp(-\frac{\pi}{2\rho^2}\int_0^1\int_0^1 \sin^2(\pi n v)\frac{1}{\sin^2(\pi u)+\sin^2(\pi v)}\,dudv). \tag{8.5}$$

The integral (8.5) can be replaced by four times the integral over the square $0 \le u, v \le 1/2$. The integral over the exterior of a small disk can be bounded independently of n. Inside a small disk

$$\frac{\sin^2(\pi u)+\sin^2(\pi v)}{\pi^2 u^2 + \pi^2 v^2} \sim 1.$$

Thus, as a rough estimate for $\exp(-\frac{1}{4}G(\epsilon_{0,n}))$ it is possible to use

$$\exp(-\frac{2}{\pi\rho^2}\int_0^{1/2} dv\{\sin^2(\pi n v)\int_0^{1/2}\frac{1}{u^2+v^2}du\})$$

and then

$$\exp(-\frac{2}{\pi\rho^2}\int_0^{1/2} dv\{\sin^2(\pi n v)(\pi/2v)\}) = \exp(-\frac{1}{\rho^2}\int_0^{\pi n/4}\frac{\sin^2(v)}{v}dv).$$

Since the integral appears to behave like $\ln(n)/2$, the term $\exp(-\frac{1}{4}G(\epsilon_{0,n}))$ behaves apart from a factor bounded away from 0 and ∞ like $n^{-\beta'}$. Thus

$$N\sum\exp(-\frac{1}{4}G(\epsilon_{0,n}))$$

behaves like $N^{2-\beta'}$ if $\beta' < 1$, like $N\ln N$ if $\beta' = 1$, and like N if $\beta' > 1$ as asserted.

The calculation at the edge is similar. The constraints are now different, the boundary conditions being free. The eigenvectors of H are now the functions

$$(ae^{\pi iux} + be^{-\pi iux})e^{2\pi ivy}/\sqrt{MN}, \qquad v = k/N, \quad 0 \le k < N,$$

but the condition on u, a, and b is that

$$\begin{gathered} a + b = ae^{\pi iu} + be^{-\pi iu}, \\ ae^{\pi iuM} + be^{-\pi iuM} = ae^{\pi iu(M+1)} + be^{-\pi iu(M+1)}. \end{gathered}$$

Thus $u = k/M$, $0 \le k < M$, and, if $u \ne 0$, we choose

$$a = \frac{c}{e^{\pi iu} - 1}, \quad b = -\frac{c}{e^{-\pi iu} - 1}, \quad c = \frac{\sqrt{2}\sin(\pi u/2)}{i},$$

so that

$$ae^{\pi iux} + be^{-\pi iux} = \sqrt{2}\cos(\pi u(x - 1/2)).$$

The eigenvectors are then normalized. If $u = 0$ take $a = b = 1/2$. The eigenvalues are $4\sin^2(\pi u/2) + 4\sin^2(\pi v)$, so that they are only small for u close to 0 and v close to 0 and 1. Thus a factor 4 that appeared in the previous calculation is replaced by 2.

The Fourier transform of $(\delta_{(1,m)} - \delta_{(1,n)})/\rho$ is, except for $u = 0$,

$$\sqrt{2}\cos(\pi u/2)(e^{2\pi imv} - e^{2\pi inv})/\rho\sqrt{MN}.$$

The square of the factor $\sqrt{2}$ adds a factor 2 restoring the factor 4. However the integral

$$\int_0^1 \frac{1}{u^2 + v^2}\,du \sim \frac{\pi}{2v}$$

is replaced by

$$\int_0^1 \frac{1}{u^2/4 + v^2}du \sim \frac{\pi}{v},$$

so that the calculations remain the same except that β' is replaced by β.

It is not without interest to reproduce these results by a formal calculation in the scaling limit. When the cylinder becomes very long the only contribution to (6.6) that is not attenuated is that for $m = n = 0$. What remains is the product over k of the two factors (6.8) and (6.9) and the second approaches 1 as q goes to 0. Thus the partition function, in its dependence on the boundary condition φ at one end, is

$$\prod_{k=1}^{\infty} \exp(-2k|a_k|^2) = \exp(-2\sum_k k|a_k|^2). \tag{8.6}$$

This defines, in an appropriate sense, a Gaussian measure on the set of periodic functions of mean 0. More to the point here, if one fixes a size N for purposes of approximation then it is appropriate to consider only functions whose Fourier expansions contain no term with wave number greater than N. On these functions (8.6) certainly defines a Gaussian measure.

On this space the analogue of S/N is

$$S_{\text{av}} = 1/2\pi \int_0^{2\pi} \exp(i\varphi(\theta)/\rho)d\theta = \int_0^1 \exp(i\sum_k a_k e^{2\pi ik\theta}/\rho)d\theta,$$

the sum running over k with $k \neq 0$ and $|k| \leq N$. Once again the mean is 0 and the scale of $S_{\rm av}$ is given by the square root of the expectation of $|S_{\rm av}|^2$. This is

$$\frac{\int_0^1 \int_0^1 \int \dots \int \prod_{k=1}^N da_k \{\exp(i \sum a_k e^{2\pi i k\theta}/\rho - i \sum a_k e^{-2\pi i k\phi}/\rho) \exp(-\sum 2k|a_k|^2)\} d\theta d\phi}{\int_0^1 \int_0^1 \int \dots \int \prod_{k=1}^N da_k \{\exp(-\sum 2k|a_k|^2) d\theta d\phi\}}$$

This quotient is

$$\int_0^1 \int_0^1 \exp(-2\sum_{k=1}^N \sin^2(2\pi k(\theta-\phi))/k\rho^2) d\theta d\phi,$$

or simplifying

$$\int_0^1 \exp(-2\sum_{k=1}^N \sin^2(2\pi k\theta)/k\rho^2) d\theta. \tag{8.7}$$

Because of the periodicity and symmetry in the integrand the integral in (8.7) can be taken from 0 to $1/4$, provided the result is multiplied by the factor 4, irrelevant to our purposes.

According to the value of β, the expression (8.7) should behave like the square of one of the expression (B.1), (B.2) or (B.3) divided by N^2. A casual argument is to set

$$a(0,\theta) = 0, \qquad a(k,\theta) = \sum_1^k 2\sin^2(2\pi l\theta) = (k+1/2) - \frac{\sin 2\pi(2k+1)\theta}{2\sin 2\pi\theta}, \quad k > 0,$$

and to sum by parts. The contributions from $a(0,\theta)$ and $a(N,\theta)$ are unimportant. Since

$$\sum_{k=1}^{N-1} \frac{k+1/2}{k(k+1)} - \ln N$$

approaches a limit we are going to get a factor $N^{-\beta}$.

A factor

$$\int_0^{1/4} \exp(-\beta \sum_{k=1}^{N-1} \frac{\sin 2\pi(2k+1)\theta}{2\sin 2\pi\theta} \frac{1}{k(k+1)}) d\theta \tag{8.8}$$

remains. The argument, apart from the factor $-\beta$, is

$$\sum_{k=1}^{N-1} \frac{\sin 2\pi(2k+1)\theta}{4\pi(k+1)\theta} \frac{2\pi\theta}{\sin 2\pi\theta} \frac{1}{k}. \tag{8.9}$$

If $2\pi/\theta \geq 1/N$, the sum can be split into two parts. In the sum over $1 < k+1 \leq 1/2\pi\theta$ the first two fractions are bounded above and below and, for small θ, are for a good many k about 1. Thus this sum is about $\ln(1/2\pi\theta)$. In the sum over $1/2\pi\theta < k+1 \leq N$ the first fraction can be replaced by $1/4\pi(k+1)\theta$ and the second by 1. Since

$$\sum_{k+1>1/2\pi\theta} \frac{1}{k(k+1)\theta} \sim 1,$$

the sum causes no problems.

The expression (8.8) is the sum of the integrals over $[0, 1/2\pi N]$ and $[1/2\pi N, 1/4]$. The second integral behaves like

$$\int_{1/2\pi N}^{1/4} \theta^{-\beta},$$

thus like a positive constant if $\beta < 1$, like $\ln N$ if $\beta = 1$ and like $N^{\beta-1}$ if $\beta > 1$. The integrand on the interval $[0, 1/2\pi N]$ is estimated as N^{β} and the integral as $N^{\beta-1}$, so that it is of the same or lower order. Since (8.8) is to be multiplied by $N^{-\beta}$, the behaviour is that expected.

The calculations in the interior are almost identical. The difference is that the pertinent measure is obtained by multiplying the weights for the partition functions on the two sides, thus by doubling the exponents in (8.6). The consequence is to replace β by β'.

Although these calculations are crude and simple, they render the statistical behaviour of the Ising model a little less puzzling – to me at least – and do suggest possibilities for further reflection.

References

[A] M. Aizenman, *The Geometry of Critical Percolation and Conformal Invariance,* to appear in Proc. of STATPHYS 19 (Xiamen 1995), World Scientific

[C1] J. Cardy, *Boundary Conditions, Fusion Rules and the Verlinde Formula,* Nucl. Phys. B324 (1989) 581-596

[C2] —, *Conformal Invariance and Statistical Mechanics,* in *Fields, Strings, and Critical Phenomena,* Les Houches (1988)

[C3] —, *Critical percolation in finite geometries,* J. Phys. A25 (1992) L201-L206

[Ch] R. Chatterjee, *Exact Partition Function and Boundary State of Critical Ising model with Boundary Magnetic Field,* Mod. Phys. Lett. A10 (1995) 973-984

[DD] B. Duplantier and F. David, *Exact Partition Functions and Correlation Functions of Multiple Hamiltonian Walks on the Manhattan Lattice,* Jour. of Stat. Phys., vol. 51 (1988) 327-434

[F] M. Fisher, *On Hearing the Shape of a Drum,* Jour. Comb. Theory, vol 1 (1966) 105-125

[G] P. Ginsbarg, *Applied Conformal Field Theory,* in *Fields, Strings, and Critical Phenomena,* Les Houches (1988)

[GJ] J. Glimm and A. Jaffe, *Quantum Physics,* Springer-Verlag (1981)

[GKP] R. L. Graham, D. E. Knuth, and O. Patashnik, *Concrete Mathematics,* Addison-Wesley 1991

[LL] R. P. Langlands and Marc-André Lafortune, *Finite Models for Percolation,* in *Contemporary Mathematics,* Amer. Math. Soc. vol. 177, (1994) 227-246

[LPPS] R. P. Langlands, Claude Pichet, Phillipe Pouliot, and Yvan Saint-Aubin, *On the Universality of Crossing Probabilities in Two-dimensional Percolation,* Jour. of Stat. Phys., vol. 67 (1992) 553-574

[LPS] R. P. Langlands, Phillipe Pouliot, and Yvan Saint-Aubin, *Conformal Invariance in Two-dimensional Percolation,* Bull. Amer. Math. Soc., (1994) 1-61

[MMS] A. Messager and S. Miracle-Sole, *Equilibrium States of the Two-dimensional Ising Model in the Two-phase Region,* Commun. Math. Phys., vol. 40 (1975) 187-196

[MW] B. M. McCoy and T. T. Wu, *The Two-dimensional Ising Model,* Harvard University Press 1973

[MS] G. Moore and Nathan Seiberg, *Classical and Quantum Conformal Field Theory,* Commun. Math. Phys., vol. 123 (1989) 177-254

[O] L. Onsager,*Crystal Statistics. I. A Two-dimensional Model with an Order-disorder Transition,* Phys. Rev. vol. 65 (1944) 117-149

[PT1] J. Palmer and C. A. Tracy, *Two-dimensional Ising Correlations: Convergence of the Scaling Limit,* Adv. in Appl. Math., vol. 2 (1981) 329-388

[PT2] J. Palmer and C. A. Tracy, *Two-dimensional Ising correlations: the SMJ Analysis,* Adv. in Appl. Math., vol. 4 (1983) 46-102

[SA] Y. Saint-Aubin, *Conformal Invariance of a Model of Percolation on Random Lattices,* **in** Stat-Phys-Taipei-1995 Conference on Nonlinear and Random Processes (Taipei, July 1995), to appear in Physica A.

[SO] R. Schor and M. O'Carroll, *The Scaling Limit and Osterwalder-Schrader Axioms for the Two-dimensional Ising Model,* Commun. Math. Phys., vol. 84 (1982) 153-170

[T] C. A. Tracy, *Asymptotics of a τ-Function Arising in the Two-dimensional Ising Model,* Commun. Math. Phys., vol. 142 (1991) 297-311

[WMTB] T. T. Wu, B. M. McCoy, C. A. Tracy, E. Barouch, *Spin-spin Correlation Functions for the Two-dimensional Ising Model: Exact Theory in the Scaling Region,* Phys. Rev. B, vol. 13 (1976) 316-374.