Toward a theory of the integer quantum Hall transition: Continuum limit of the Chalker−Coddington model

Martin R. Zirnbauer
Institut für Theoretische Physik, Universität zu Köln, Germany

An N-channel generalization of the network model of Chalker and Coddington is considered. The model for $N=1$ is known to describe the critical behavior at the plateau transition in systems exhibiting the integer quantum Hall effect. Using a recently discovered equality of integrals, the network model is transformed into a lattice field theory defined over Efetov's σ model space with unitary symmetry. The transformation is exact for all N, no saddle-point approximation is made, and no massive modes have to be eliminated. The naive continuum limit of the lattice theory is shown to be a supersymmetric version of Pruisken's nonlinear σ model with couplings $\sigma_{xx}=N/4$ and $\sigma_{xy}=N/2$ at the symmetric point. It follows that the model for $N=2$, which describes a spin degenerate Landau level and the random flux problem, is noncritical. On the basis of symmetry considerations and inspection of the Hamiltonian limit, a modified network model is formulated, which still lies in the quantum Hall universality class. The prospects for deformation to a Yang−Baxter integrable vertex model are briefly discussed. © *1997 American Institute of Physics.* [S0022-2488(97)01204-8]

I. INTRODUCTION

The single-electron states of a two-dimensional disordered electron gas in a strong magnetic field are localized except at the energies of the Landau band centers. As the Fermi energy approaches such a band center, a critical phenomenon takes place: the localization length diverges and the Hall conductance jumps from one plateau to the next. This phase transition, which belongs to the general class of Anderson metal−insulator transitions, has been seen in several experiments and a substantial amount of data on its critical behavior is available from a number of numerical simulations (in the absence of electron−electron interactions), see Ref. 1 and references therein. Unfortunately, in spite of considerable efforts expended over the last decade, our analytical understanding of the plateau-to-plateau transition is still rather poor. It is expected that the critical behavior is described by some nonunitary conformal field theory, but this theory has not yet been identified.

There exist two opposite limits from which the transition in the noninteracting system can be approached theoretically. The first limit is that of a slowly varying random potential with a correlation length l_c much larger than the magnetic length l_B. In this limit, the electron's motion can be described in semiclassical terms.[2] More precisely, the motion separates into a rapid cyclotron motion superimposed on a slow guiding center drift along spatially localized equipotential lines. As the Fermi energy approaches the center of a Landau band, a percolating path develops and a localization−delocalization transition takes place. Close to the transition, the quantum mechanical possibility for an electron to tunnel from an equipotential to a neighboring one is a relevant perturbation. The essential features of this quantum percolation transition were cast into a random network model by Chalker and Coddington.[3] In their model an electron acquires random U(1) phases while moving along the directed bonds of a square network, and is scattered to the right or left every time it passes through a node of the network. The model will be reviewed in more detail below. Suffice it to say here that the model has been studied by numerical simulation but has, in its original, spatially isotropic formulation, defied analytical treatment up to now,

although a certain amount of analytical insight has come from consideration of its anisotropic[4] and weak disorder[5] versions.

The opposite limit is $l_c \ll l_B$. Historically, this was the *first* limit to be studied analytically, the reason being that it is this limit that was amenable to the field theoretic machinery developed at the beginning of the 1980's by Wegner,[6] Efetov[7] and others. Starting from a Gaussian white-noise potential ($l_c = 0$) Pruisken[8] used the replica trick to set up a generating functional for the disorder averaged retarded and advanced Green's functions of the single-electron Hamiltonian. After a Hubbard–Stratonovitch transformation to matrix-valued fields, he made a saddle-point approximation, valid in the limit of a high Landau level. This was followed by a gradient expansion leading to a $U(n_+ + n_-)/U(n_+) \times U(n_-)$ nonlinear σ model with a vanishing number of replicas, $n_+ = n_- = 0$, and a parity-violating topological term due to the strong magnetic field. The coupling constants of the model, σ_{xx} and σ_{xy}, were identified[9] with the physical conductivities of the $2d$ disordered electron gas. The topological coupling $\theta = 2\pi\sigma_{xy}$, by its very nature, has no effect on the equations of motion of the classical field theory, but *does* change both the Hamiltonian and the symplectic structure and, consequently, the commutator of the quantum theory. It was argued[10] that the nonlinear σ model, while generically being massive (i.e., having a finite correlation length) in two dimensions, becomes critical at $\theta = \pi$. The vanishing of the mass gap corresponds to the appearance of delocalized states at the center of a Landau band. Thus, Pruisken's model provided the right kind of scenario in which to develop a scaling theory of the plateau-to-plateau transition. A supersymmetric version of the model first appeared in Ref. [11].

Undeniably, Pruisken's nonlinear σ model has served as a great inspiration to theory. One of its early successes was the $\sigma_{xx} - \sigma_{xy}$ flow that was conjectured from it[12,10] and later verified by numerical and real experiments; see Ref. 1 and references therein. In spite of this, Pruisken's model or, rather, its promoters have been criticized; see Ref. 13 for a summary. For one thing, the model has never yielded any quantitative results for the critical behavior at the transition, and much less has it been solved. (The same statement applies to the general class of nonlinear σ models with a topological term. None of these has ever been solved, at least not directly.) For another, even the derivation of the model is vulnerable to criticism: the validity of the saddle-point approximation that is made to eliminate the so-called massive modes, requires $\sigma_{xx} \gg 1$. Although this inequality is satisfied for the bare (or SCBA) value of σ_{xx} in the limit of a high Landau level, the renormalized theory at $\sigma_{xy} = 1/2$ is expected to have a σ_{xx} of order unity or less. The cure proposed by Pruisken was to *assume* the renormalizability of his model, and appeal to the RG flow to take the coupling constant σ_{xx} from large to small values. However, such an assumption needs to be justified and, in fact, is not acceptable by current field theoretic knowledge, for Pruisken's model apparently lacks the conformal structure that is required of a fixed point theory with a continuous symmetry in two dimensions. (In other words, the model, while definitely being critical at $\sigma_{xy} = 1/2$, does not possess the conservation laws expected of an infrared stable fixed point.)

Two advances will be made in the present paper. The first is to establish a very close connection between Pruisken's nonlinear σ model and the network model of Chalker and Coddington at criticality. We will show that the latter can be viewed as a lattice discretization of the former or, conversely, taking the continuum limit of the network model yields the σ model. A more detailed outline is the following. We start out by reviewing the supersymmetric version of Pruisken's model and the Chalker–Coddington network model in Secs. II and III. Then, in Sec. IV, the network model is reformulated as a lattice-regularized field theory defined over Efetov's super-symmetric nonlinear σ model space with unitary symmetry. This reformulation is exact. In contrast with the conventional method due to Wegner, Efetov and others, no saddle-point approximation is needed to eliminate the massive modes. Moreover, Sec. V shows that not only is the supersymmetric reformulation of the network model defined over the same field space, but it also has the same global symmetries as Pruisken's model. At the critical point of the network model, where the correlation (or localization) length diverges, the long wave length physics of the supersymmetric lattice theory is expected to be described by a continuum field theory. The symmetries

dictate that this continuum theory be Pruisken's model at the critical coupling $\sigma_{xy} = 1/2$. As is shown in Sec. VII A, the other coupling constant, σ_{xx}, equals 1/4. The numerical value of σ_{xy} is checked in Sec. VII B, by evaluating the lattice action on a smooth field configuration with nonzero topological charge. Section VII C extends these results to an N-channel network model with random U(N) matrices on links.

What we learn from all this is that, although historically Pruisken's model was first obtained from the Gaussian white noise limit $l_c \ll l_B$, in a low Landau level it is actually more closely related to the opposite limit $l_c \gg l_B$, since it is the latter that provides the microscopic basis for the network model.

The mathematical basis underlying the exact transformation from the N-channel network model to the supersymmetric lattice field theory is quite natural and simple, and is briefly sketched as follows. (Readers who are not interested in mathematical structures may want to skip this paragraph.) For a pair of positive integers n, N consider the tensor product $\mathbb{C}^n \otimes \mathbb{C}^N$, on which the group GL($nN$) acts by linear transformations. Span the corresponding Lie algebra gl(nN) by bilinears $\{\bar{c}_A^i c_B^j\}_{A,B=1,\ldots,N}^{i,j=1,\cdots,n}$ in fermionic creators $\bar{c}$ and annihilators c, which act in a Fock space with vacuum $|0\rangle$. There exist two natural subalgebras: gl(n), generated by $\Sigma_i \bar{c}_A^i c_B^i$, and gl($N$), generated by $\Sigma_A \bar{c}_A^i c_A^j$. Now put $n = n_- + n_+$ and fill the "negative energy states" to form $|\text{vac}\rangle = \Pi_{i=1}^N \Pi_{A=1}^{n_-} \bar{c}_A^i |0\rangle$. The particle–hole coherent states[14] that are generated by the action of GL(n) on $|\text{vac}\rangle$, are holomorphic sections of a line bundle associated to the homogeneous space $G/H := \text{GL}(n)/\text{GL}(n_+) \times \text{GL}(n_-)$ by the Slater-determinant representation of GL(n_-) on $|\text{vac}\rangle$. They are parameterized by a complex $n_+ \times n_-$ matrix Z with adjoint $Z^\dagger$. By combining the closure relation for particle–hole coherent states with a few elementary properties of Fermi-coherent (or Grassmann-coherent) states, one can prove the following equality of integrals:

$$\int_{\text{U}(N)} dU \prod_{a=1}^{n_+} \text{Det}(1 - e^{+i\varphi_{+a}} U) \prod_{b=1}^{n_-} \text{Det}(1 - e^{-i\varphi_{-b}} U^\dagger)$$

$$= \int d\mu(Z, Z^\dagger) \text{Det}^{-N}(1 + Z^\dagger Z) \text{Det}^N(1 + Z^\dagger e^{+i\varphi_+} Z e^{-i\varphi_-}),$$

where $\varphi_\pm = \text{diag}(\varphi_{\pm 1}, \ldots, \varphi_{\pm n_\pm})$ are diagonal matrices with real entries, dU is the Haar measure of a U(N) subgroup of GL(N) and $d\mu(Z, Z^\dagger)$ expresses the U(n)-invariant measure of a compact symmetric space U(n)/U(n_+)×U(n_-) contained in the complex homogeneous space G/H. This integral identity forms the mathematical basis of our formalism. Its supersymmetric extension[15] permits us to carry out the disorder average over the random U(N) matrices placed on the links of the network, at the expense of introducing an integration over fields Z taking values in a symmetric superspace.

Another issue addressed in this paper is the question whether the supersymmetric formulation of the network model offers the possibility for an exact analytical solution. In Sec. VI I reveal that the model resembles what is called a *vertex model* in statistical physics, in the sense that the Boltzmann weight is a product of factors, one for each node, or vertex. The weight associated with a single vertex is called the R-matrix. The symmetries of the R-matrix are investigated in Sec. VIII. It is eventually found that it can be interpreted as a map $\mathscr{R}: V \otimes V \to V \otimes V$, where V is an irreducible lowest-weight module for the Lie superalgebra gl(n,n), and $n = n_+ + n_- = 1 + 1$ for the case of one retarded and one advanced Green's function. This looks interesting as one may hope that $\mathscr{R}$ can be deformed to an R-matrix that solves the quantum Yang–Baxter equation underlying the integrability of two-dimensional vertex models. One would then have the possibility of an analytical and exact computation of critical properties. Unfortunately, the specific choice of local directions for the single-particle motion on the network, shown in Fig. 2 below, turns out to be *incompatible* with the standard schemes[16] for constructing solutions of the quantum Yang–Baxter equation. The reason is that the R-matrix of a Yang–Baxter solvable model always trans-

fers from one side of the vertex to the other, whereas the Chalker–Coddington vertex maps the horizontal degrees of freedom into the vertical ones, or vice versa. Thus the Chalker–Coddington model in its original form does not fit into the canonical framework of the theory of integrable systems, and I am not aware of any method to make analytical progress with it.

However, this is not yet the end of the story. Additional insight can be gained by considering the anisotropic limit[4] of the Chalker–Coddington model. This limit and its relation to Pruisken's model are reviewed in Sec. IX. Based on it, in Sec. X a modified version of the isotropic one-channel network model is proposed, which differs from the original one in two respects. First, the direction of the single-particle motion does not alternate constantly between being horizontal and vertical. Instead, an electron may either pass straight through a node (with *no* change of direction), or else be scattered either to the right or to left. Analysing such a model by the mapping onto Pruisken's nonlinear σ model, we find that it is likely *not* to be in the quantum Hall universality class, but in a massive (Haldane type) phase. Therefore, a second modification is proposed, which is to add a second channel of propagation on *half* the links, say the horizontal ones; see Fig. 5(b) below. By the mapping onto Pruisken's model, the doubly modified network model is expected to be critical in a range of values of the parameter characterizing the scattering at the nodes. Moreover, by the changed transfer dynamics, the standard schemes for constructing solutions of the quantum Yang–Baxter equation are no longer ruled out. I hope to elaborate on this theme in a future publication.

II. SUPERSYMMETRIC FORMULATION OF PRUISKEN'S MODEL (DEFINITIONS)

The original formulation[8] of Pruisken's nonlinear σ model relied on the use of the replica trick. When applied to phenomena that are nonperturbative in the disorder strength, the replica trick is not mathematically sound but has been demonstrated to lead to incorrect results, at least in some instances.[17] (The analytic continuation to a vanishing number of replicas is not unique in general.) Fortunately, we can avoid the replica trick by using an alternative, supersymmetric formalism,[7] which is on firm mathematical ground. The purpose of this section is to briefly review the supersymmetric formulation[11] of Pruisken's model in a language that is well suited for what will follow below. For simplicity, only the model pertaining to one retarded and one advanced Green's function is treated. A more detailed discussion of the model can be found in Ref. 13.

To define Pruisken's nonlinear σ model we first specify its field space, as follows. Consider a pair $Z, \widetilde{Z}$ of 2×2 complex supermatrices,

$$Z = \begin{pmatrix} Z_{BB} & Z_{BF} \\ Z_{FB} & Z_{FF} \end{pmatrix}, \quad \widetilde{Z} = \begin{pmatrix} \widetilde{Z}_{BB} & \widetilde{Z}_{BF} \\ \widetilde{Z}_{FB} & \widetilde{Z}_{FF} \end{pmatrix},$$

where the subscripts B and F stand for Bosonic and Fermionic, and let

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in GL(2|2)$$

act on these by

$$Z \mapsto g \cdot Z = (AZ + B)(CZ + D)^{-1}, \quad \widetilde{Z} \mapsto g \cdot \widetilde{Z} = (C + D\widetilde{Z})(A + B\widetilde{Z})^{-1}.$$

Following Ref. 13 one identifies the pair $Z, \widetilde{Z}$ as a set of coordinates for the complex coset space $G/H := Gl(2|2)/GL(1|1) \times GL(1|1)$, where the denominator is the subgroup generated by the block-diagonal matrices

$$h = \begin{pmatrix} A & 0 \\ 0 & D \end{pmatrix}.$$

With this identification, the action $Z \mapsto g \cdot Z$ and $\widetilde{Z} \mapsto g \cdot \widetilde{Z}$ coincides[13] with the natural action of G on G/H by left translation. The supermatrices Z and $\widetilde{Z}$ then are viewed as left-translates of the origin in G/H, by writing $Z \equiv g \cdot 0 = BD^{-1}$ and $\widetilde{Z} \equiv g \cdot \widetilde{0} = CA^{-1}$.

The coset space G/H, being a homogeneous space, admits only one (up to multiplication by a constant) rank-two supersymmetric tensor that is invariant under the action of $G = \mathrm{GL}(2|2)$. In the coordinates $Z, \widetilde{Z}$ it is given by

$$\hat{g} = \mathrm{STr}(1 - Z\widetilde{Z})^{-1}\mathrm{d}Z(1 - \widetilde{Z}Z)^{-1}\mathrm{d}\widetilde{Z},$$

where STr denotes the supertrace. The G-invariant superintegration measure that derives from $\hat{g}$ is denoted $D(Z, \widetilde{Z})$ and has the local coordinate expression[18]

$$D(Z, \widetilde{Z}) = \mathrm{d}Z_{\mathrm{BB}} \wedge \mathrm{d}\widetilde{Z}_{\mathrm{BB}} \wedge \mathrm{d}Z_{\mathrm{FF}} \wedge \mathrm{d}\widetilde{Z}_{\mathrm{FF}} \frac{\partial^4}{\partial Z_{\mathrm{BF}} \partial \widetilde{Z}_{\mathrm{BF}} \partial Z_{\mathrm{FB}} \partial \widetilde{Z}_{\mathrm{FB}}}.$$

(Note that superintegration measures, also called integral superforms or Berezin forms, generically suffer from a coordinate ambiguity, or anomaly; see Ref. 19. A general procedure by which to define $D(Z, \widetilde{Z})$ *globally* is described in Sec. II A of Ref. 20.) The integration domain for the bosonic variables is fixed by the conditions

$$\widetilde{Z}_{\mathrm{FF}} = -\overline{Z}_{\mathrm{FF}}, \quad \widetilde{Z}_{\mathrm{BB}} = +\overline{Z}_{\mathrm{BB}} \quad \text{and} \quad |Z_{\mathrm{BB}}|^2 < 1,$$

where the bar denotes complex conjugation. These conditions select a submanifold $M_{\mathrm{B}} \times M_{\mathrm{F}}$ of G/H,

$$M_{\mathrm{B}} = \mathrm{U}(1,1)/\mathrm{U}(1) \times \mathrm{U}(1) \simeq \mathrm{H}^2, \quad M_{\mathrm{F}} = \mathrm{U}(2)/\mathrm{U}(1) \times \mathrm{U}(1) \simeq \mathrm{S}^2,$$

on which the metric $\hat{g}$ is Riemann. The variables Z_{FF} and Z_{BB} can be shown[13] to have a meaning as complex stereographic coordinates on the two-sphere S^2 and two-hyperboloid H^2, respectively. The triple $(G/H, M_{\mathrm{B}} \times M_{\mathrm{F}}, \hat{g})$ is a Riemannian symmetric superspace,[20] and is called Efetov's σ model space with unitary symmetry. In most of the existing literature, this space is parameterized by a 4×4 supermatrix Q, introduced below.

After these preparations, Pruisken's nonlinear σ model (or, rather, the supersymmetric version thereof) is defined by the functional integral,

$$\langle \bullet \rangle = \int \mathscr{D}(Z, \widetilde{Z}) \bullet \exp - S_{\mathrm{cont}}[Z, \widetilde{Z}],$$

where $S_{\mathrm{cont}}[Z, \widetilde{Z}] = \int L \, d^2 r$ is obtained by integrating the following Lagrangian:

$$L = \sigma_{xx}^{(0)}(\mathscr{L}_{xx} + \mathscr{L}_{yy}) + \sigma_{xy}^{(0)}(\mathscr{L}_{xy} - \mathscr{L}_{yx}),$$

$$\mathscr{L}_{\mu\nu} = \mathrm{STr}(1 - Z\widetilde{Z})^{-1}\partial_\mu Z(1 - \widetilde{Z}Z)^{-1}\partial_\nu \widetilde{Z}, \tag{1}$$

and the functional integration measure is

$$\mathscr{D}(Z, \widetilde{Z}) = \prod_{\mathbf{r} = (x,y)} D(Z(\mathbf{r}), \widetilde{Z}(\mathbf{r})).$$

By construction, this field theory is invariant under global $\mathrm{Gl}(2|2)$ transformations $Z(\mathbf{r}) \mapsto g \cdot Z(\mathbf{r})$ and $\widetilde{Z}(\mathbf{r}) \mapsto g \cdot \widetilde{Z}(\mathbf{r})$ (each of the terms $\mathscr{L}_{\mu\nu}$ is). The partition function $\mathscr{Z} = \langle 1 \rangle$ equals unity by supersymmetry. To obtain physical observables, one adds sources and takes functional derivatives, as usual.

For some purposes it is useful to switch to a coordinate-independent language by introducing a 4×4 supermatrix field Q by

$$Q = \begin{pmatrix} 1 & Z \\ \widetilde{Z} & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & Z \\ \widetilde{Z} & 1 \end{pmatrix}^{-1},$$

in terms of which the Lagrangian takes the familiar form[8]

$$\mathscr{L}_{xx} + \mathscr{L}_{yy} = \mathrm{STr}(\partial_x Q \partial_x Q + \partial_y Q \partial_y Q)/8 =: L_0,$$

$$\mathscr{L}_{xy} - \mathscr{L}_{yx} = \mathrm{STr}(\partial_x Q \partial_y Q - \partial_y Q \partial_x Q)Q/8 =: L_{\mathrm{top}}.$$

The global action of $\mathrm{GL}(2|2)$ on Q is $Q(\mathbf{r}) \mapsto g Q(\mathbf{r}) g^{-1}$. On a configuration manifold C without boundary the integral of L_{top} is quantized:[10,18] $\int_C L_{\mathrm{top}} d^2 r = 2\pi i n$, and the integer n is called the ''winding number'' or ''topological charge.'' L_0 and L_{top} are invariant under rotations in the xy plane, and these are the only invariants that can be formed from $\mathscr{L}_{\mu\nu}$. Note that L_0 is real-valued, whereas L_{top} is purely imaginary. The coupling constants $\sigma_{xx}^{(0)}$ and $\sigma_{xy}^{(0)}$ have been interpreted[9] as the longitudinal and Hall conductivities of the 2d electron gas.

Let me mention in passing that the terms in the Lagrangian can be written as

$$L_0 = \left(\frac{\partial^2}{\partial x \partial x'} + \frac{\partial^2}{\partial y \partial y'} \right) \ln \mathrm{SDet}(1 - Z(x,y)\widetilde{Z}(x',y'))^{-1} \big|_{x=x', y=y'},$$

$$L_{\mathrm{top}} = \left(\frac{\partial^2}{\partial x \partial y'} - \frac{\partial^2}{\partial y \partial x'} \right) \ln \mathrm{SDet}(1 - Z(x,y)\widetilde{Z}(x',y'))^{-1} \big|_{x=x', y=y'},$$

(2)

where SDet denotes the superdeterminant. The function $\ln \mathrm{SDet}(1 - Z\widetilde{Z})$ is called the Kähler potential of the Kähler supermanifold G/H parameterized by the complex coordinates Z and $\widetilde{Z}$.

III. THE CHALKER–CODDINGTON MODEL

The Chalker–Coddington model was formulated[3] for the purpose of describing the plateau-to-plateau transition in systems exhibiting the integer quantum Hall effect. Its microscopic foundation as a model for the motion of a single 2d electron subject to a smooth random potential and a strong magnetic field, is explained in the original paper.

The model is defined on a finite or infinite square lattice forming a network of nodes and directed links; see Fig. 1. A ''wave function'' of the model is a collection of complex amplitudes $\{\psi(l)\}$, one for each link l of the network. The dynamics is governed not by a Hamiltonian but by a unitary operator U, called the one-step time evolution operator. U acts on wave functions, and is given by a sequence of two distinct operations. The first one is stochastic and, in a microscopic picture, accounts for the random phase acquired by the electron's guiding center while drifting along equipotentials between saddle points of the random potential. In formal language, the stochastic phases are encapsulated in a unitary operator U_0 that is diagonal in the basis provided by the links: $U_0(l',l) = \delta(l',l) \exp i\varphi(l)$, where $\varphi(l)$ are uncorrelated random variables with a uniform distribution on the interval $[0, 2\pi]$.

The second process built into the one-step time evolution operator is deterministic (i.e., nonrandom) and accounts for the quantum mechanical possibility for an electron to tunnel across a saddle point. This is modeled via the nodes of the network. One imagines that the guiding center
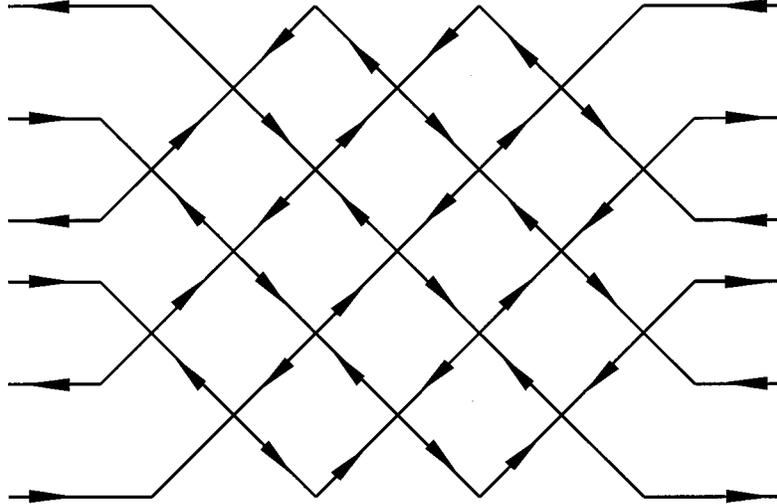
FIG. 1. The network model of Chalker and Coddington in a quasi 1d geometry. A single electron propagates ballistically along the links and is scattered at the nodes of the network.

drift motion of the electron follows the direction indicated by the arrows in Fig. 1. An electron incident on a node can be scattered either to the left or to the right, corresponding to the two possibilities of continuing on its way along the same equipotential, or tunneling to a neighboring one. It cannot be backscattered, and it cannot pass straight through a node. The probabilities for scattering to the right or left are denoted p and $1-p$. The probability *amplitudes* for scattering at the nodes then have magnitude $\sqrt{p}$ and $\sqrt{1-p}$, respectively. They are taken to be real, but they cannot all be chosen positive as this would violate unitarity. Various choices are possible. For definiteness we take the amplitudes to be negative for right-up to left-up and for right-down to left-down turns (see Fig. 1) and positive otherwise. (Which choice is made actually turns out to be immaterial. All we need is that some choice consistent with unitarity exists.) All this fixes a unitary operator U_1 with matrix elements $U_1(l',l)$ that vanish unless the scattering process $l \to l'$ is allowed by the (one-step) dynamics, in which case $U_1(l',l) = \pm\sqrt{p}$ or $U_1(l',l) = \pm\sqrt{1-p}$, as specified. The full one-step time evolution operator is the product $U = U_1 U_0$.

For the purpose of doing numerical simulations, one usually takes the network to be a long strip (quasi-1d geometry) with periodic boundary conditions for the short direction to minimize finite size effects. By computing the exponential growth of the transfer matrix for the strip and averaging over many realizations of the disorder embodied by U_0, one extracts a Lyapunov exponent or inverse localization length. For $p = 1/2$ the localization length is found to grow linearly with the width of the strip, indicating a critical point.

The existence of critical behavior at $p = 1/2$ can be anticipated by the following argument.[3] Consider the Chalker–Coddington model with the open boundary conditions of Fig. 1. For $p = 0$ (left turns only) all electron states encircle elementary squares of the network in the counterclockwise direction, and thus are strongly localized. The same statement applies for $p = 1$ (right turns only), except that now the orientation of the motion around squares is clockwise, and an *edge state* at the boundary of the strip appears, as is seen by inspection of Fig. 1. The appearance of an edge state implies that somewhere in the interval $p \in [0,1]$ a delocalized state must form. For symmetry reasons, this is expected to happen at the left–right symmetric point $p = 1/2$.

The method to be introduced in Sec. IV is quite capable of dealing with a quasi-1d geometry and its boundary conditions. However, for maximal simplicity we will consider below a somewhat different setup, where periodic boundary conditions are imposed in both directions, i.e., the net-
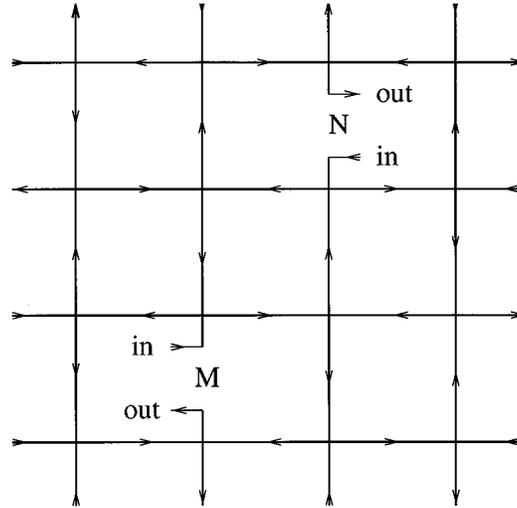
FIG. 2. Chalker–Coddington network with two interior contacts M and N. The corresponding links l_M and l_N are connected to one incoming and one outgoing channel each.

work is placed on a torus $S^1 \times S^1$. We envisage making a conductance measurement between two *interior* contacts M and N as shown in Fig. 2. These contacts are ''point'' contacts in the sense that they each attach only to a *single* link, l_M and l_N. For most purposes it is helpful to imagine that each link with a contact is replaced by two links, an ''in-link'' feeding the network from a reservoir, and an ''out-link'' draining the network through an outgoing channel leading back to the reservoir. Outgoing-wave boundary conditions are imposed on out-links, i.e. probability flux that is scattered into such a link by the action of U_1, must exit and never returns to the network. The conductance pertaining to two interior contacts M and N can be computed from the Landauer–Büttiker formula, $g_{MN} = |S_{MN}|^2$, where S_{MN} is the S-matrix element relating an incoming state on the in-link of l_N to an outgoing wave amplitude on the out-link of l_M. This S-matrix element is determined by the solution of Schrödinger's equation $U\psi = e^{i\alpha}\psi$ with incoming-wave boundary conditions at N. The eigenphase $e^{i\alpha}$ may be gauged away by a global shift of the random U(1) factors on links, $U_0(l) \mapsto e^{i\alpha} U_0(l)$. It is convenient to implement the loss of probability flux through out-links by the modification $U_0(l) = \exp i\varphi(l) \to U_0(l^{\text{out}}) \equiv 0$ for every link l with a contact. Then, solving the equation $\psi = U\psi$ by iteration, we easily see that the d.c. conductance equals

$$g_{MN} = |S_{MN}|^2 = |\langle l_M^{\text{out}}| T | l_N^{\text{in}} \rangle|^2,$$

where T is the operator

$$T = U_1 + U_1 U_0 U_1 + U_1 U_0 U_1 U_0 U_1 + \cdots = (1 - U_1 U_0)^{-1} U_1. \tag{3}$$

Note that owing to $U_0(l_M^{\text{out}}) = U_0(l_N^{\text{out}}) = 0$, the eigenvalues of $U_1 U_0$ lie *inside* the unit circle (rather than on the unit circle) in $\mathbb{C}$, so the inverse $(1 - U_1 U_0)^{-1}$ is well-defined. Without changing the matrix element $|S_{MN}|^2$, we set $U_0(l_M^{\text{in}}) = U_0(l_N^{\text{in}}) = 0$.

IV. REFORMULATION BY SUPERSYMMETRY

In this section we will show how to map the problem of calculating the disorder averaged conductance $\langle g_{MN} \rangle$ on an equivalent lattice field theory. From $g_{MN} = |\langle l_M^{\text{out}}| T | l_N^{\text{in}} \rangle|^2$, $T = (1 - U_1 U_0)^{-1} U_1$, and $T^\dagger = (1 - U_1^\dagger U_0^\dagger)^{-1} U_1^\dagger$, the mean conductance $\langle g_{MN} \rangle$ equals the disorder average of the expression

$$\sum_{l_m l_n} \langle l_N^{\text{in}}|(1-U_1^\dagger U_0^\dagger)^{-1}|l_m\rangle U_1^\dagger(l_m,l_M^{\text{out}})\langle l_M^{\text{out}}|(1-U_1 U_0)^{-1}|l_n\rangle U_1(l_n,l_N^{\text{in}}).$$

This disorder average can be processed by a variant of Efetov's supersymmetry method introduced in Ref. 15, as follows. The method starts by expressing the product of matrix elements as a Gaussian superintegral:

$$\langle l_N^{\text{in}}|(1-U_1^\dagger U_0^\dagger)^{-1}|l_m\rangle\langle l_M^{\text{out}}|(1-U_1 U_0)^{-1}|l_n\rangle$$

$$= \int \prod_l D(\psi(l),\overline{\psi}(l))\,\psi_{-\text{B}}(l_N^{\text{in}})\,\overline{\psi}_{-\text{B}}(l_m)\,\psi_{+\text{B}}(l_M^{\text{out}})\,\overline{\psi}_{+\text{B}}(l_n)$$

$$\times \exp(-\overline{\psi}_{+\sigma}(l')[\,\delta(l',l)-U_1(l',l)U_0(l)]\psi_{+\sigma}(l)$$

$$-\overline{\psi}_{-\tau}(l')[\,\delta(l',l)-\overline{U}_1(l,l')\overline{U}_0(l)]\psi_{-\tau}(l)),$$

where $\psi_{\pm\sigma}$ are the components of a complex superfield $l\mapsto\psi(l)$. The index $\sigma=\text{B}$ or F distinguishes between Bosonic and Fermionic components, and $\pm$ relates to retarded (T) and advanced $(T^\dagger)$ Green's functions. The bar denotes complex conjugation. $D(\psi,\overline{\psi})$ denotes the ''flat'' superintegration measure, i.e., is given by the product of the differentials of the bosonic components times the product of partial derivatives with respect to the fermionic ones. The summation convention is used here and throughout the paper.

For the following step, it is notationally convenient to absorb U_1 temporarily by setting $\hat{\psi}_{+\sigma}(l)=\overline{\psi}_{+\sigma}(l')U_1(l',l)$ and $\hat{\psi}_{-\tau}(l)=\overline{\psi}_{-\tau}(l')\overline{U}_1(l,l')$. We will now trade the average over random phases $\int\prod_l dU_0(l)=\int\prod_l d\varphi(l)/2\pi$ for an integral over Efetov's σ model space with unitary symmetry, reviewed in Sec. II. This is done by a kind of Hubbard–Stratonovitch transformation,

$$\int \prod_l{}' dU_0(l)\exp(\hat{\psi}_{+\sigma}(l)U_0(l)\psi_{+\sigma}(l)+\hat{\psi}_{-\tau}(l)\overline{U}_0(l)\psi_{-\tau}(l))$$

$$= \int \prod_l{}' D\mu(Z(l),\widetilde{Z}(l))\exp(\hat{\psi}_{+\sigma}(l)Z_{\sigma\tau}(l)\psi_{-\tau}(l)+\hat{\psi}_{-\tau}(l)\widetilde{Z}_{\tau\sigma}(l)\psi_{+\sigma}(l)), \tag{4}$$

which is a special case of a more general result proved in Ref. 15. The 2×2 supermatrix valued fields $l\mapsto Z(l)$ and $l\mapsto\widetilde{Z}(l)$ are lattice discretizations of the continuous fields of Pruisken's model. The superintegration measure,

$$D\mu(Z,\widetilde{Z}):=\text{const}\times D(Z,\widetilde{Z})\text{SDet}(1-\widetilde{Z}Z),$$

is normalized by the condition $\int D\mu(Z,\widetilde{Z})=1$. Because of the boundary condition at the contacts, $U_0(l_M)=U_0(l_N)=0$ (in- and out-links), the product over links on both sides of (4) excludes l_M and l_N. An alternative scheme of implementing the boundary condition is to let the product run over *all* links, including the boundary ones, and set

$$Z(l_M)=\widetilde{Z}(l_M)=Z(l_N)=\widetilde{Z}(l_N)=0. \tag{5}$$

We adopt this scheme. By using (4) to deal with the disorder average of the product of matrix elements, we obtain

$$\langle\langle l_N^{\text{in}}|(1-U_1^\dagger U_0^\dagger)^{-1}|l_m\rangle\langle l_M^{\text{out}}|(1-U_1 U_0)^{-1}|l_n\rangle\rangle$$

$$= \int \prod_l D\mu(Z(l),\widetilde{Z}(l)) \int \prod_l D(\psi(l),\overline{\psi}(l))$$

$$\times \psi_{-\text{B}}(l_N^{\text{in}})\overline{\psi}_{-\text{B}}(l_m)\psi_{+\text{B}}(l_M^{\text{out}})\overline{\psi}_{+\text{B}}(l_n)$$

$$\times \exp(-\overline{\psi}_{+\sigma}(l)\psi_{+\sigma}(l)+\overline{\psi}_{+\sigma}(l')U_1(l',l)Z_{\sigma\tau}(l)\psi_{-\tau}(l)$$

$$-\overline{\psi}_{-\tau}(l)\psi_{-\tau}(l)+\overline{\psi}_{-\tau}(l')\overline{U}_1(l,l')\widetilde{Z}_{\tau\sigma}(l)\psi_{+\sigma}(l)).$$

The final step is to carry out the Gaussian integral over ψ and $\overline{\psi}$, which gives

$$\langle g_{MN}\rangle = \int \prod_l D\mu(Z(l),\widetilde{Z}(l))\text{SDet}_{\mathscr{H}}(1-Z_U\widetilde{Z})^{-1}[Z_U(1-\widetilde{Z}Z_U)^{-1}]_{\text{BB}}(l_M^{\text{out}},l_M^{\text{out}})$$

$$\times [\widetilde{Z}_{U^\dagger}(1-Z\widetilde{Z}_{U^\dagger})^{-1}]_{\text{BB}}(l_N^{\text{in}},l_N^{\text{in}}),$$

by an elementary calculation. Here Z_U denotes the supermatrix field Z evolved forward in time by one action of the deterministic scattering operator U_1, i.e., $Z_U = U_1 Z U_1^\dagger$ or, in more explicit notation,

$$(U_1 Z U_1^\dagger)_{\sigma\tau}(l,l') = U_1(l,l'')Z_{\sigma\tau}(l'')\overline{U}_1(l',l'').$$

Similarly, $\widetilde{Z}_{U^\dagger}$ is the supermatrix field $\widetilde{Z}$ evolved backward in time by one step: $\widetilde{Z}_{U^\dagger} = U_1^\dagger \widetilde{Z} U_1$. The subscript $\mathscr{H}$ on SDet indicates that the superdeterminant runs over both superspace and the Hilbert space of wave functions supported on links.

   In summary, we have shown that the average conductance pertaining to two interior contacts M and N can be computed as a two-point function $\langle g_{MN}\rangle = \langle \mathscr{O}(M)\widetilde{\mathscr{O}}(N)\rangle$ of operators

$$\mathscr{O}(M) = [Z_U(1-\widetilde{Z}Z_U)^{-1}](l_M^{\text{out}},l_M^{\text{out}}),$$

$$\widetilde{\mathscr{O}}(N) = [\widetilde{Z}_{U^\dagger}(1-Z\widetilde{Z}_{U^\dagger})^{-1}](l_N^{\text{in}},l_N^{\text{in}}),$$

in a lattice field theory,

$$\langle\bullet\rangle = \int \prod_l D(Z(l),\widetilde{Z}(l))\bullet\exp-S_{\text{latt}}[Z,\widetilde{Z}],$$

with lattice action

$$S_{\text{latt}}[Z,\widetilde{Z}] = \ln\text{SDet}_{\mathscr{H}}(1-Z_U\widetilde{Z}) - \ln\text{SDet}_{\mathscr{H}}(1-Z\widetilde{Z}). \tag{6}$$

The second term in S_{latt} originates from the measure $D\mu(Z,\widetilde{Z}) = D(Z,\widetilde{Z})\text{SDet}(1-Z\widetilde{Z})$. Note that the functional integral for the two-point function $\langle \mathscr{O}(M)\widetilde{\mathscr{O}}(N)\rangle$ is regularized by the boundary condition (5).

   What we have done is an *exact rewriting* of the original problem. It might superficially look as though we have made the problem more complicated by transforming from U(1) phase integrals to an integral over supermatrix fields Z and $\widetilde{Z}$, but this is not so. The point is that while the U(1) phases $U_0(l)$ *fluctuate independently*, the dominant contributions to the integral over the supermatrix field come from *slowly varying* field configurations. The latter property is best seen by combining terms to rewrite the lattice action as follows:

$$S_{\text{latt}}[Z,\widetilde{Z}] = \ln \text{SDet}_{\mathscr{H}}(1-(1-\widetilde{Z}Z)^{-1}\widetilde{Z}(Z_U-Z)).$$

The right-hand side vanishes for constant fields ($Z_U = Z$) and is small for slowly varying ones. This will allow us to take a continuum limit and describe the low energy or long wave length physics of the Chalker–Coddington network model by a continuum field theory. In view of (2) and (6), it should not come as a surprise that this field theory will turn out to be Efetov's nonlinear σ model with unitary symmetry, augmented by Pruisken's topological term.

The formula we have derived for $\langle g_{MN} \rangle$ illustrates how the transport coefficients of the network model can be expressed as correlations functions of a supersymmetric lattice field theory. In the remainder of the paper we will not discuss the specific correlator $\langle g_{MN} \rangle$ any further, but will concentrate on the general structure of the theory.

V. GLOBAL SYMMETRY UNDER GL(2|2)

To get ready for taking the continuum limit, we will now elucidate the symmetries of the lattice field theory. We are going to show that it is invariant under global transformations,

$$Z(l) \mapsto g \cdot Z(l) = (AZ(l)+B)(CZ(l)+D)^{-1},$$

$$\widetilde{Z}(l) \mapsto g \cdot \widetilde{Z}(l) = (C+D\widetilde{Z}(l))(A+B\widetilde{Z}(l))^{-1},$$

for

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \text{GL}(2|2).$$

The integration measure $D(Z(l),\widetilde{Z}(l))$ has this invariance *by definition*. To see that the lattice action (6) is invariant, we first transform the factor $\text{SDet}(1-Z(l)\widetilde{Z}(l))$ for a single link l, temporarily dropping the argument l for notational simplicity:

$$\text{SDet}(1-(g \cdot Z)(g \cdot \widetilde{Z})) = \text{SDet}(1-(AZ+B)(CZ+D)^{-1}(C+D\widetilde{Z})(A+B\widetilde{Z})^{-1})$$

$$= \text{SDet}(1-(Z+A^{-1}B)(1+D^{-1}CZ)^{-1}(\widetilde{Z}+D^{-1}C)(1+A^{-1}B\widetilde{Z})^{-1}).$$

This expression is further processed by setting $A^{-1}B =: X$ and $D^{-1}C =: \widetilde{X}$ and using the identity

$$1-(Z+X)(1+\widetilde{X}Z)^{-1}(\widetilde{Z}+\widetilde{X})(1+X\widetilde{Z})^{-1} = (1-X\widetilde{X})(1+Z\widetilde{X})^{-1}(1-Z\widetilde{Z})(1+X\widetilde{Z})^{-1},$$

which follows from elementary algebra. We then get

$$\text{SDet}(1-(g \cdot Z)(g \cdot \widetilde{Z})) = \text{SDet}((1-X\widetilde{X})(1+Z\widetilde{X})^{-1}(1-Z\widetilde{Z})(1+X\widetilde{Z})^{-1}).$$

The ''dynamical'' factor $\text{SDet}_{\mathscr{H}}(1-Z_U\widetilde{Z})^{-1}$ is transformed in an identical fashion. Because the transformation is global and U_1 acts as the identity in superspace, we have

$$(g \cdot Z)_U = U_1(g \cdot Z)U_1^{-1} = g \cdot (U_1ZU_1^{-1}) = g \cdot Z_U,$$

i.e., the actions of g and U_1 commute. We thus obtain

$$\text{SDet}_{\mathscr{H}}(1-(g \cdot Z)_U(g \cdot Z))^{-1} = \text{SDet}_{\mathscr{H}}(1-(g \cdot Z_U)(g \cdot \widetilde{Z}))^{-1}$$

$$= \text{SDet}_{\mathscr{H}}((1-X\widetilde{X})(1+Z_U\widetilde{X})^{-1}(1-Z_U\widetilde{Z})(1+X\widetilde{Z})^{-1}).$$

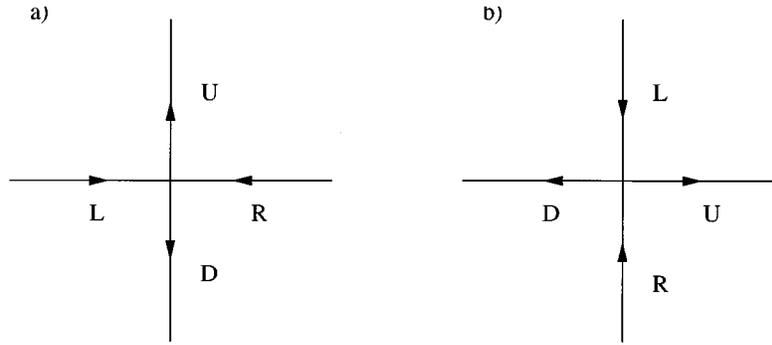By combining factors and using the multiplicativity of the superdeterminant, we arrive at

FIG. 3. The two types of elementary vertex of the Chalker–Coddington model. Type (b) is obtained from type (a) by a rotation through $\pi/2$.

$$S_{\text{latt}}[g \cdot Z, g \cdot \widetilde{Z}] = S_{\text{latt}}[Z, \widetilde{Z}] - \ln \text{SDet}_{\mathcal{H}}(1 + Z\widetilde{X})^{-1} + \ln \text{SDet}_{\mathcal{H}}(1 + Z_U\widetilde{X})^{-1}.$$

The last two terms on the right-hand side cancel each other by

$$\text{SDet}_{\mathcal{H}}(1 + Z_U\widetilde{X}) = \text{SDet}_{\mathcal{H}}(1 + U_1 Z U_1^{-1}\widetilde{X}) = \text{SDet}_{\mathcal{H}}(1 + Z\widetilde{X}),$$

since $\widetilde{X}$ is constant on links and therefore commutes with U_1 (i.e., $U_1^{-1}\widetilde{X}U_1 = \widetilde{X}$). This establishes the global invariance $S_{\text{latt}}[g \cdot Z, g \cdot \widetilde{Z}] = S_{\text{latt}}[Z, \widetilde{Z}]$.

The global GL(2|2) symmetry of S_{latt} is broken by the boundary condition (5), of course. Such a symmetry breaking is needed for the mathematical consistency of the formulation, since the integral over the $q = 0$, or zero momentum, mode of the noncompact bosonic sector would otherwise be divergent.

In Sec. IV we saw that the supersymmetric reformulation of the Chalker–Coddington model is defined over the same field space as Pruisken's model. What we have just learned is that these models not only share the field space, but also have the *same internal symmetries*. This is already a strong indication of their equivalence at the critical point $p = 1/2$, resp. $\sigma_{xy}^{(0)} = 1/2$.

VI. REFORMULATION AS A VERTEX MODEL

Further analysis is facilitated by the observation that the ''Boltzmann weight'' of the lattice field theory,

$$W[Z, \widetilde{Z}] = \text{SDet}_{\mathcal{H}}(1 - \widetilde{Z}Z_U)^{-1}\text{SDet}_{\mathcal{H}}(1 - \widetilde{Z}Z),$$

factors to a large extent. This factorization is seen as follows. Consider the expression $\widetilde{Z}(l'')U_1(l'',l')Z(l')U_1^{-1}(l',l)$ and take l to be the link $l = D$, say, emanating in the downward direction from the node shown in Fig. 3(a). The matrix element $U_1^{-1}(l',D)$ is nonzero only if l' is one of the two links, $l' = L$ or $l' = R$ in the figure, that flow into the same node. In either case, the matrix element $U_1(l'',l')$ leads to the final state l'' being one of only two links, namely $l'' = l = D$ (the one we started from), or else $l'' = U$, the link emanating from the node in the upward direction. Similarly, if the initial state is $l = U$, the final state is $l'' = U$ or $l'' = D$. These statements remain true (mutatis mutandi) if the node of Fig. 3(a) is replaced by any other node of the network. (Half of the nodes of the network are obviously equivalent to the one considered. The other half of them has links emanating in the horizontal direction and nodes flowing into it in the vertical direction; see Fig. 3(b). These nodes become equivalent to the one of Fig. 3(a) after a rotation by ninety degrees.)

Thus, $\widetilde{Z}U_1ZU_1^{-1}$ viewed as an operator in the space of states supported on links, connects only *pairs of links*, namely the two links that emanate from one and the same node. This means that the Boltzmann weight W is a product of factors, one for each node:

$$W[Z,\widetilde{Z}] = \prod_{\text{nodes } n} \mathscr{R}(\widetilde{Z}(U_n),\widetilde{Z}(D_n),Z(L_n),Z(R_n)), \tag{7}$$

where the assignment of link labels U,D,L,R is defined by Fig. 3. From the definition of U_1 in Sec. III, the weight for a single node works out to be

$$\mathscr{R}(\widetilde{Z}(U),\widetilde{Z}(D),Z(L),Z(R))$$

$$= \text{SDet}([1-\widetilde{Z}(U)Z(U)][1-\widetilde{Z}(D)Z(D)][1-\widetilde{Z}(L)Z(L)][1-\widetilde{Z}(R)Z(R)])^{1/2}$$

$$\times \text{SDet}^{-1}\begin{pmatrix} 1-p\widetilde{Z}(U)Z(R)-(1-p)\widetilde{Z}(U)Z(L) & \sqrt{p(1-p)}\widetilde{Z}(U)[Z(L)-Z(R)] \\ \sqrt{p(1-p)}\widetilde{Z}(D)[Z(L)-Z(R)] & 1-p\widetilde{Z}(D)Z(L)-(1-p)\widetilde{Z}(D)Z(R) \end{pmatrix}.$$
$$\tag{8}$$

A model of the kind (7) is called a *vertex model* in statistical mechanics, and $\mathscr{R}$ is called the R-matrix. The global symmetries of this R-matrix will be investigated in Sec. VIII. In Sec. X we will show how to pass from the integration over the fields $Z,\widetilde{Z}$ to a summation over superspin degrees of freedom.

VII. CONTINUUM LIMIT

For values of the parameter p close to 1/2, where the Chalker–Coddington model undergoes a quantum percolation transition, the correlation (or localization) length is very large, and we expect the lattice functional integral to be dominated by fields that vary slowly. Our goal in this section is to extract from (6) the continuum field theory governing these slowly varying modes.

Recall the following facts: (i) Pruisken's model and the supersymmetric (SUSY) reformulation of the Chalker–Coddington model are defined over the very same complex field space, $G/H = \text{GL}(2|2)/\text{GL}(1|1)\times\text{GL}(1|1)$. (ii) Both the action functional of Pruisken's model, S_{cont}, and the lattice action of the SUSY reformulated Chalker–Coddington model, S_{latt}, are invariant under global $\text{GL}(2|2)$ transformations. (iii) The metric tensor of the field space G/H,

$$\hat{g} = \text{STr}(1-Z\widetilde{Z})^{-1}\text{d}Z(1-\widetilde{Z}Z)^{-1}\text{d}\widetilde{Z} = -\text{STr}(\text{d}Q)^2/8,$$

is invariant under $\text{GL}(2|2)$ [i.e., under $Q\mapsto gQg^{-1}$ with $g\in\text{GL}(2|2)$], and this is the *only* second-rank supersymmetric tensor on the homogeneous space G/H that has this invariance. When combined with a standard field theoretic power counting argument, these facts lead to the conclusion that the continuum limit of the lattice action S_{latt} must be a linear combination of the four terms $\mathscr{L}_{\mu\nu}$ ($\mu,\nu=x,y$) in (1), which are induced from $\hat{g}$.

A further constraint comes from the spatial symmetries. At large scales (i.e., in the infrared limit), where details of the network structure are washed out, the Chalker–Coddington model acquires an invariance under rotations in space. From the four terms $\mathscr{L}_{\mu\nu}$ one can form only two linear combinations that are rotationally invariant. These are precisely the ones that figure in the expression (1) for Pruisken's Lagrangian (Pruisken *et al.*[10] arrived at them by essentially the same argument), namely $\mathscr{L}_{xx}+\mathscr{L}_{yy}$ and $\mathscr{L}_{xy}-\mathscr{L}_{yx}$. We thus conclude that the continuum limit of the SUSY reformulated Chalker–Coddington model at criticality ($p=1/2$) is Pruisken's model at the critical topological coupling $\theta=\pi=2\pi\sigma_{xy}^{(0)}$, and some as yet unknown coupling $\sigma_{xx}^{(0)}$:

$$S_{\text{latt}}[\text{smooth fields}] = S_{\text{cont}}\big|_{\sigma_{xx}^{(0)}=?,\,\sigma_{xy}^{(0)}=1/2}.$$

The only thing that remains to be done, then, is to work out the numerical value of the coupling constant $\sigma_{xx}^{(0)}$. Although it is not too difficult to do this calculation for a whole range of values of the parameter p, we restrict ourselves to the critical point $p=1/2$, for simplicity. In the following subsection we will show that

$$\sigma_{xx}^{(0)} = 1/4 \quad (\text{for } p=1/2).$$

The small value of $\sigma_{xx}^{(0)}=1/4$ means that the theory is at *strong coupling*. (The weak-coupling limit, where Pruisken's derivation is valid, is $\sigma_{xx}^{(0)} \gg 1$.) The superscript (0) alerts us to the fact that this is a ''bare'' value, obtained by a naive continuum limit, i.e. without taking into account possible renormalizations coming from modes with short wave lengths.

In the following we pay no attention to the boundary conditions and assume the network to be infinitely extended.

A. The coupling constant $\sigma_{xx}^{(0)}$

To show that $\sigma_{xx}^{(0)}=1/4$ for the Chalker–Coddington model at $p=1/2$, we first expand S_{latt} to quadratic order in Z and $\widetilde{Z}$ and then take the long wave length limit. The quadratic part of S_{latt} is conveniently obtained by expanding the expression (10) for the R-matrix:

$$-2\ln \mathscr{R}(\widetilde{Z}(U),\widetilde{Z}(D),Z(L),Z(R)) = \text{STr}(\widetilde{Z}(U)[Z(U)-Z(L)]+[\widetilde{Z}(R)-\widetilde{Z}(U)]Z(R)+\widetilde{Z}(D)$$
$$\times[Z(D)-Z(R)]+[\widetilde{Z}(L)-\widetilde{Z}(D)]Z(L)) + \mathscr{O}(Z^2\widetilde{Z}^2).$$

Now recall the meaning of the labels U,D,L,R defined by Fig. 3(a), and let $(n_x,n_y) \in \mathbb{Z}^2$ be the Cartesian coordinates of the node in that figure. Then, if a is the lattice constant of the network,

$$Z(L)=Z(an_x-a/2,an_y), \quad Z(D)=Z(an_x,an_y-a/2),$$

$$Z(R)=Z(an_x+a/2,an_y), \quad Z(U)=Z(an_x,an_y+a/2).$$

Similar expressions hold for the other type of node, shown in Fig. 3(b) (rotate by $\pi/2$). We sum $\ln \mathscr{R}$ over all nodes of the lattice and go to momentum space with wave vector $\mathbf{k}=(k_x,k_y)$. The quadratic part S_2 of S_{latt} is then obtained as

$$S_2 = \frac{1}{2}\sum_{\mathbf{k}} t_{\mathbf{k}}\text{STr}\widetilde{Z}(\mathbf{k})Z(-\mathbf{k}),$$

where

$$t_{\mathbf{k}} = 4 - e^{i(k_x+k_y)a/2} - e^{-i(k_x+k_y)a/2} - e^{i(k_x-k_y)a/2} - e^{-i(k_x-k_y)a/2}.$$

On taking the continuum limit $a \to 0$, and replacing the sum over momenta by an integral, we get

$$S_2 = \frac{1}{4}\int \frac{d^2k}{(2\pi)^2}\mathbf{k}^2 \text{ STr } \widetilde{Z}(\mathbf{k})Z(-\mathbf{k}).$$

Matching this expression to the quadratic part of S_{cont}, Eq. (1), gives $\sigma_{xx}^{(0)}=1/4$, as was claimed above.

B. The coupling constant $\sigma_{xy}^{(0)}$

The critical points of the Chalker–Coddington model and Pruisken's model are at $p = 1/2$ and $\sigma_{xy}^{(0)} = 1/2$ (mod 1), respectively. Therefore, as was argued earlier, there cannot be any doubt that the continuum limit of the lattice action (6) at $p = 1/2$ is Pruisken's action with topological coupling $\sigma_{xy}^{(0)} = 1/2$. Nevertheless, it is both reassuring and instructive to check this by direct calculation, which is what we do next.

In Sec. VII A the value of $\sigma_{xx}^{(0)}$ was found by looking at the quadratic part of the action. The topological coupling $\sigma_{xy}^{(0)}$, by its very nature, evades any such attempt at perturbative calculation. To extract it from (6) a different, nonperturbative scheme must be used. A direct approach would be to perform a gradient expansion around a topologically nontrivial background. For reasons that will be explained at the end of Sec. VIII, this is not easy to do. Here we will follow a different procedure, which is to evaluate S_{latt} on the lattice discretization $Z^{(m)}$ of some smooth field configuration with topological charge $m \neq 0$.

Given the relation $\mathrm{Im} S_{\text{contl}}[Z^{(m)}, \widetilde{Z}^{(m)}] = 2\pi m \sigma_{xy}^{(0)}$ and the requirement $S_{\text{cont}} = S_{\text{latt}}$ for smooth fields, the topological coupling is determined by

$$\mathrm{Im}\, S_{\text{latt}}[Z^{(m)}, \widetilde{Z}^{(m)}] = 2\pi m \sigma_{xy}^{(0)}.$$

To calculate $\sigma_{xy}^{(0)}$ from this equation, it is easiest to consider fields that have topological charge $m = 1$ and are of the special form

$$Z^{(1)} = \begin{pmatrix} 0 & 0 \\ 0 & f \end{pmatrix}, \quad \widetilde{Z}^{(1)} = \begin{pmatrix} 0 & 0 \\ 0 & -\overline{f} \end{pmatrix}.$$

Here the components Z_{BB}, Z_{BF}, and Z_{FB}, which are topologically trivial, have been set to zero and only Z_{FF} has been retained. With this choice, the formula for $\sigma_{xy}^{(0)}$ reduces to

$$\sigma_{xy}^{(0)} = \frac{1}{2\pi}\, \mathrm{Im}\, S_{\text{latt}}[Z^{(1)}, \widetilde{Z}^{(1)}]$$

$$= -\frac{1}{2\pi}\, \mathrm{Im}\, \ln \mathrm{Det}_{\mathscr{H}}(1 + \overline{f} U_1 f U_1^{-1})$$

$$= -\frac{1}{2\pi} \sum_{\text{nodes}\ n}\, \mathrm{Im}\, \ln \mathscr{R}_{\text{FF}}(\overline{f}(U_n), \overline{f}(D_n), f(L_n), f(R_n)),$$

where $\mathscr{R}_{\text{FF}}$ is the R-matrix in the FF sector:

$$\mathscr{R}_{\text{FF}}(\overline{f}(U_n), \overline{f}(D_n), f(L_n), f(R_n)) = \mathrm{Det}\begin{pmatrix} 1 + \overline{f}(U)[f(L) + f(R)]/2 & \overline{f}(U)[f(L) - f(R)]/2 \\ \overline{f}(D)[f(L) - f(R)]/2 & 1 + \overline{f}(D)[f(L) + f(R)]/2 \end{pmatrix}$$

$$= 1 + \frac{1}{2}[\overline{f}(U) + \overline{f}(D)][f(L) + f(R)] + \overline{f}(U)\overline{f}(D)f(L)f(R).$$

Now consider in the 2d plane with coordinates x and y the smooth field configuration

$$f(x, y) = \left(\frac{x - y}{A_-} + i\frac{x + y}{A_+} - z_0\right)^{-1} \quad (A_\pm \in \mathbb{R}),$$

which interpolates between $f = \infty$ (corresponding to the south pole on the two-sphere $M_{\text{F}} = S^2$) at

$$x + iy = \frac{1-i}{2} A_- \, \mathrm{Re}\, z_0 + \frac{1+i}{2} A_+ \, \mathrm{Im}\, z_0, \qquad (9)$$

and $f = 0$ (the north pole on S^2) at infinity, and has topological charge $m = 1$, as can easily be checked. (Note that f is *not* a one-instanton configuration, unless $A_- = A_+$.) The parameters $A_\pm$ set the size in $x \pm y$ of the region where the topological density L_{top} for f is appreciably different from zero.

For this choice of f, we now work out the R-matrix at the node in Fig. 3(a) with coordinates (n_x, n_y). Setting $\alpha_\pm = a/(2A_\pm)$, and $\zeta = (n_x - n_y)a/A_- + i(n_x + n_y)a/A_+ - z_0$, we have

$$\overline{f}(U) = (\overline{\zeta} - \alpha_- - i\alpha_+)^{-1}, \quad \overline{f}(D) = (\overline{\zeta} + \alpha_- + i\alpha_+)^{-1},$$

$$f(L) = (\zeta - \alpha_- - i\alpha_+)^{-1}, \quad f(R) = (\zeta + \alpha_- + i\alpha_+)^{-1}.$$

The resulting value for the R-matrix is

$$\mathscr{R}_{\mathrm{FF}} = 1 + \frac{2|\zeta|^2 + 1}{[\zeta^2 - (\alpha_- + i\alpha_+)^2][\overline{\zeta}^2 - (\alpha_- + i\alpha_+)^2]}.$$

In order for the field configuration to be smooth on the lattice, the parameters $\alpha_\pm$ must be small. This allows us to Taylor expand with respect to one of these, say α_+. Taking also the imaginary part of the logarithm, we obtain

$$\mathrm{Im}\, \ln \mathscr{R}_{\mathrm{FF}} = 4\alpha_+ \alpha_- \rho(1+\rho)^{-1} \, \mathrm{Re}(\zeta^2 - \alpha_-^2)^{-1} + \mathcal{O}(\alpha_+^2),$$

where $1 + \rho = \mathscr{R}_{\mathrm{FF}}|_{\alpha_+ = 0}$.

In the next step we sum the contributions from all nodes $(n_x + k, n_y + k)$ with $k \in \mathbb{Z}$. The smallness of $\alpha_+ \ll 1$ allows us to convert the expression for $\Sigma \, \mathrm{Im}\, \ln \mathscr{R}_{\mathrm{FF}}$ into an integral. Decomposing ζ into real and imaginary parts by $\zeta = \xi + i\eta$,

$$\xi = 2\alpha_-(n_x - n_y) - \mathrm{Re}\, z_0, \quad \eta = 2\alpha_+(n_x + n_y) - \mathrm{Im}\, z_0,$$

we arrive at

$$\sum_k \, \mathrm{Im}\, \ln \mathscr{R}_{\mathrm{FF}}\big|_{(n_x + k, n_y + k)} \xrightarrow{\alpha_+ \to 0} \alpha_- \int_{\mathbb{R}} d\eta \, \frac{\rho}{1+\rho} \, \mathrm{Re}((\xi + i\eta)^2 - \alpha_-^2)^{-1}.$$

This integral is easily evaluated by closing the contour and applying the method of residues. One finds

$$\alpha_- \int_{\mathbb{R}} d\eta \, \frac{\rho}{1+\rho} \, \mathrm{Re}((\xi + i\eta)^2 - \alpha_-^2)^{-1} = \begin{cases} -\pi, & \text{if } |\xi| < \alpha_-; \\ 0, & \text{otherwise.} \end{cases}$$

What is the geometric interpretation of this result? Let us agree that the word ''vertex'' here means a node taken together with its four links, severed at half the distance to neighboring nodes. The set of vertices with node coordinates $(n_x + k, n_y + k)$ ($k \in \mathbb{Z}$) sweep out a diagonal strip of the 2d network. The above result means that $\Sigma_k \, \mathrm{Im}\, \ln \mathscr{R}_{\mathrm{FF}}$ vanishes if the center (9) of the topological excitation lies outside the diagonal strip swept out by the vertices $(n_x + k, n_y + k)$ (with variable $k \in \mathbb{Z}$), and equals $-\pi$ when it lies inside.

The above calculation applies to the type of node shown in Fig. 3(a), which may be characterized by the condition $n_x + n_y \in 2\mathbb{Z}$, say. Doing the calculation for the other type of node

$[n_x + n_y \in 2\mathbb{Z}+1$, or Fig. 3(b)] gives the same result except for a change of sign $-\pi \to \pi$. Now observe that the diagonal strips swept out by the vertices $(n+k, -n+k)$ and $(n+k, -n+1+k)$ $(k, n \in \mathbb{Z})$ cover the plane completely without overlapping. Therefore,

$$\exp i \, \text{Im} \, S_{\text{latt}}[Z^{(1)}, \widetilde{Z}^{(1)}] = \exp \sum_{\text{all nodes}} -i \, \text{Im} \, \ln \mathscr{R}_{\text{FF}} = e^{\pm i\pi} = -1,$$

independent of the location of the topological singularity. By $\sigma_{xy}^{(0)} = (2\pi)^{-1} \, \text{Im} \, S_{\text{latt}}[Z^{(1)}, \widetilde{Z}^{(1)}]$, this proves $\sigma_{xy}^{(0)} = 1/2$ (mod 1).

C. Generalization to *N* channels

Lee and Chalker[21] introduced a generalization of the network model that has *two* channels per link. The one-step time evolution operator of that model, $U^{(2)}$, is again a product of factors: $U^{(2)} = U_1^{(2)} U_0^{(2)}$. The second factor is diagonal on links and associates with each link l a 2×2 matrix $U_0(l)$ drawn at random from the unitary group in two-dimensional channel space, U(2). The first factor describes the deterministic scattering at the nodes. Depending on the choice made for this factor, the two-channel model applies to a spin degenerate Landau level or electrons in a random magnetic field (the so-called random flux problem).

In this subsection we consider an *N*-channel generalization of the one- and two-channel network models, where random U(N) matrices are placed on the links. The one-step time evolution operator is written $U^{(N)} = U_1^{(N)} U_0^{(N)}$. The two factors describe the deterministic *N*-channel scattering at the nodes, and the random U(N) directed propagation along links, respectively.

The case of N channels per link can be treated by a slight extension of our field theoretic formalism. Such an extension is possible since the basic formula (4) was shown in Ref. 15 to generalize from U(1) to U(N), as follows:

$$\int_{U(N)} dU \, \exp(\overline{\psi}_{+\sigma}^i U^{ij} \psi_{+\sigma}^j + \overline{\psi}_{-\tau}^j \overline{U}^{ij} \psi_{-\tau}^i) = \int_{\text{Efetov}} D\mu_N(Z, \widetilde{Z}) \exp(\overline{\psi}_{+\sigma}^i Z_{\sigma\tau} \psi_{-\tau}^i + \overline{\psi}_{-\tau}^j \widetilde{Z}_{\tau\sigma} \psi_{+\sigma}^j),$$

where $U \equiv U_0(l)$, and the link label l was suppressed for clarity. The right-hand side differs from that of (4) only by the channel index $i = 1,...,N$ attached to $\psi, \overline{\psi}$, and the different form of the weight function in the superintegration measure:

$$D\mu_N(Z, \widetilde{Z}) := D(Z, \widetilde{Z})\text{SDet}(1 - Z\widetilde{Z})^N.$$

Using the above generalization of (4), we can reformulate the *N*-channel model as a supersymmetric lattice field theory with action

$$S_{\text{latt}}^{(N)}[Z, \widetilde{Z}] = \ln \text{SDet}_{\mathscr{H} \otimes \mathbb{C}^N}(1 - U_1^{(N)} Z U_1^{(N)\dagger} \widetilde{Z}) - N \ln \text{SDet}_{\mathscr{H}}(1 - Z\widetilde{Z}).$$

The derivation exactly parallels that of Sec. IV. The first superdeterminant runs over the tensor product of superspace with link space ($\mathscr{H}$) and channel space ($\mathbb{C}^N$).

If we choose the deterministic scattering $U_1^{(N)}$ to be of the special form $U_1^{(N)} = U_1^{(N=1)} \otimes 1_N$, i.e. $U_1^{(N)}$ acts as the identity in channel space, $S_{\text{latt}}^{(N)}$ is simply a multiple of the action of the one-channel model:

$$S_{\text{latt}}^{(N)}[Z, \widetilde{Z}] = N \times S_{\text{latt}}^{(N=1)}[Z, \widetilde{Z}].$$

Therefore, from Secs. VII A and VII B the coupling constants of the corresponding continuum field theory are

$$\sigma_{xx}^{(0)} = N/4, \quad \sigma_{xy}^{(0)} = N/2.$$

The choice $U_1^{(N)} = U_1^{(N=1)} \otimes 1_N$ for $N=2$ is appropriate[21] for the spin degenerate Landau level at the center of the band, and for the random flux problem at the symmetric point. Because Pruisken's model at $\sigma_{xy}^{(0)} = 2/2 \pmod 1 = 0$ is known to be a massive theory, our result for $\sigma_{xy}^{(0)}$ confirms the proposition of Ref. 21 that neither of these systems is critical. Rather, they are in a (Haldane type) massive phase,[4] corresponding to localization of all states.

VIII. SYMMETRIES OF THE R-MATRIX

The basic building block of the SUSY reformulated Chalker–Coddington model is its R-matrix. For the case $N=1$, which is to be analysed here in more detail, this building block was given in (8). To gain a deeper understanding of the model and possible variations thereof, we are now going to investigate the symmetries of that R-matrix. Most of the effort will be expended on rewriting the expression (8) in a different form, so as to make those symmetries more evident.

We start by undoing the $\psi, \overline{\psi}$ integration to write the R-matrix as

$$\mathscr{R} = \prod_{I,O} \mathrm{SDet}(1 - \widetilde{Z}(O)Z(O))^{1/2} \, \mathrm{SDet}(1 - \widetilde{Z}(I)Z(I))^{1/2} \int D(\psi, \overline{\psi}) \exp(- \overline{\psi}_{+\sigma}(O)\psi_{+\sigma}(O)$$

$$+ \overline{\psi}_{+\sigma}(O)U_1(O,I)Z_{\sigma\tau}(I)\psi_{-\tau}(I) - \overline{\psi}_{-\tau}(I)\psi_{-\tau}(I) + \overline{\psi}_{-\tau}(I)U_1^{-1}(I,O)\widetilde{Z}_{\tau\sigma}(O)\psi_{+\sigma}(O)).$$

Here we have introduced the labels $I \in \{i1,i2\} := \{L,R\}$ and $O \in \{o1,o2\} := \{U,D\}$. The notation is motivated by observing that, according to the direction of motion indicated by the arrows in Fig. 3, the links L and R are *incoming* states for the scattering at the nodes, while the links U and D are *outgoing* states.

To proceed, we need to recall briefly various mathematical structures that were developed in detail in the appendices of Refs. 15 and 22. First of all, we introduce Fock operators c and $\overline{c}$, which are quantum counterparts of the classical variables ψ and $\overline{\psi}$. Let $b_{\pm}^{\dagger}, b_{\pm}$ and $f_{\pm}^{\dagger}, f_{\pm}$ be canonical boson and fermion creation and annihilation operators, and set

$$c_{+F} = f_+, \quad c_{+B} = b_+, \quad c_{-F} = f_-^{\dagger}, \quad c_{-B} = b_-^{\dagger},$$

$$\overline{c}_{+F} = f_+^{\dagger}, \quad \overline{c}_{+B} = b_+^{\dagger}, \quad \overline{c}_{-F} = f_-, \quad \overline{c}_{-B} = -b_-.$$

The operators c and $\overline{c}$ are canonical pairs, with graded (or super) commutation relations

$$[c_X, \overline{c}_Y] := c_X \overline{c}_Y - (-1)^{|X||Y|} \overline{c}_Y c_X = \delta_{XY},$$

where $|X| = 0$ if $X = \pm B$ and $|X| = 1$ if $X = \pm F$. They act in a Bose–Fermi Fock space with vacuum

$$c_{+B}|0\rangle = c_{+F}|0\rangle = \overline{c}_{-B}|0\rangle = \overline{c}_{-F}|0\rangle = 0. \tag{10}$$

The graded commutation relations are invariant under canonical transformations, $\overline{c}_X \mapsto T_g \overline{c}_X T_g^{-1} = \overline{c}_Y g_{YX}$ and $c_X \mapsto T_g c_X T_g^{-1} = (g^{-1})_{XY} c_Y$, where

$$g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{GL}(2|2) \mapsto T_g := \exp(\overline{c}_X (\ln g)_{XY} c_Y)$$

defines a representation of $\mathrm{GL}(2|2)$ on Fock space.

Consider now the subspace, V, selected by the condition $\overline{c}_X c_X = 0$ (summation convention!) or, equivalently, by

$$b^\dagger_+ b_+ + f^\dagger_+ f_+ = b^\dagger_- b_- + f^\dagger_- f_- .$$

The resulting constraint on the Bose–Fermi occupation numbers is $n_{b+} + n_{f+} = n_{b-} + n_{f-}$. By using this equation to eliminate n_{b-}, say, we can characterize the states of V by a triplet of integers (n_{b+}, n_{f+}, n_{f-}), where n_{f+} and n_{f-} take values from the set $\{0,1\}$, and $n_{b+} = 0,1,2,\dots,\infty$.

Let P_V be the operator that projects Fock space onto the subspace V. As was shown in Ref. 15, this projector has a resolution,

$$P_V = \int D(Z,\widetilde{Z}) |Z\rangle\langle Z|,$$

by generalized coherent states,

$$|Z\rangle = \exp(\overline{c}_{+\sigma} Z_{\sigma\tau} c_{-\tau}) |0\rangle \mathrm{SDet}(1 - Z\widetilde{Z})^{1/2},$$

$$\langle Z| = \mathrm{SDet}(1 - Z\widetilde{Z})^{1/2} \langle 0| \exp(-\overline{c}_{-\tau} \widetilde{Z}_{\tau\sigma} c_{+\sigma}).$$

By simple manipulations exploiting the standard properties of coherent states, one can verify the following equality:

$$\prod_I \mathrm{SDet}(1 - \widetilde{Z}(I) Z(I))^{1/2} \times \exp(\overline{\psi}_{+\sigma}(O) U_1(O,I) Z_{\sigma\tau}(I) \psi_{-\tau}(I))$$

$$= \langle 0| \exp(\overline{\psi}_{+\sigma}(o\mu) U_1(o\mu, i\nu) c_{+\sigma}(\nu) - \overline{c}_{-\tau}(\nu) \psi_{-\tau}(i\nu)) |Z(i1) \otimes Z(i2)\rangle$$

$$= \langle 0| \exp(\overline{\psi}_{+\sigma}(o\mu) c_{+\sigma}(\mu) - \overline{c}_{-\tau}(\nu) \psi_{-\tau}(i\nu))$$

$$\times \exp(\overline{c}_{+\sigma}(\mu)(\ln \hat{U}_1)(\mu,\nu) c_{+\sigma}(\nu)) |Z(i1) \otimes Z(i2)\rangle.$$

Here $\mu, \nu \in \{1,2\}$, and $\hat{U}_1$ is defined by identifying initial and final channels, i.e., $\hat{U}_1(\mu,\nu) = U_1(o\mu, i\nu)$. Similarly,

$$\prod_O \mathrm{SDet}(1 - \widetilde{Z}(O) Z(O))^{1/2} \times \exp(\overline{\psi}_{-\tau}(I) U_1^{-1}(I,O) \widetilde{Z}_{\tau\sigma}(O) \psi_{+\sigma}(O))$$

$$= \langle Z(o1) \otimes Z(o2)| \exp(\overline{\psi}_{-\tau}(i\nu) U_1^{-1}(i\nu, o\mu) c_{-\tau}(\mu) + \overline{c}_{+\sigma}(\mu) \psi_{+\sigma}(o\mu)) |0\rangle$$

$$= \langle Z(o1) \otimes Z(o2)| \exp(\overline{c}_{-\tau}(\nu)(\ln\hat{U}_1)(\nu,\mu) c_{-\tau}(\mu))$$

$$\times \exp(\overline{c}_{+\sigma}(\mu) \psi_{+\sigma}(o\mu) + \overline{\psi}_{-\tau}(i\nu) c_{-\tau}(\nu)) |0\rangle.$$

The variables $\psi, \overline{\psi}$ have now served their purpose and we integrate them out, by using the closure relation for Bose–Fermi coherent states:

$$\mathrm{id} = \int D(\psi,\overline{\psi}) \exp(-\overline{\psi}_{+\sigma}(\mu) \psi_{+\sigma}(\mu) - \overline{\psi}_{-\tau}(\nu) \psi_{-\tau}(\nu)) \exp(\overline{c}_{+\sigma}(\mu) \psi_{+\sigma}(\mu)$$

$$+ \overline{\psi}_{-\tau}(\nu) c_{-\tau}(\nu)) |0\rangle\langle 0| \exp(\overline{\psi}_{+\sigma}(\mu) c_{+\sigma}(\mu) - \overline{c}_{-\tau}(\nu) \psi_{-\tau}(\nu)).$$

All this results in the following formula for the R-matrix:

$$\mathscr{R}(\widetilde{Z}(U), \widetilde{Z}(D), Z(L), Z(R)) = \langle Z(U) \otimes Z(D)| \hat{\mathscr{R}} |Z(L) \otimes Z(R)\rangle,$$

where the operator $\hat{\mathscr{R}}$ is expressed by

$$\hat{\mathcal{R}} = \exp(\overline{c}_X(\mu)(\ln \hat{U}_1)(\mu,\nu)c_X(\nu)). \tag{11}$$

The advantage of this reformulation is that the invariance of $\hat{\mathcal{R}}$ under GL(2|2) transformations,

$$\overline{c}_X(\mu) \mapsto T_g \overline{c}_X(\mu) T_g^{-1} = \overline{c}_Y(\mu) g_{YX},$$

$$c_X(\mu) \mapsto T_g c_X(\mu) T_g^{-1} = (g^{-1})_{XY} c_Y(\mu),$$

is obvious whereas previously, in Sec. VI, we had to work quite hard to establish the global GL(2|2) invariance.

Although $\hat{\mathcal{R}}$ commutes with T_g, its matrix elements are *not* invariant:

$$\mathcal{R}(\widetilde{Z}(U),\widetilde{Z}(D),Z(L),Z(R)) \neq \mathcal{R}(g \cdot \widetilde{Z}(U), g \cdot \widetilde{Z}(D), g \cdot Z(L), g \cdot Z(R)).$$

It is worth spending a little effort to explain how that comes about. As will be seen, the reason is that, since the Fock space vacuum is not a scalar with respect to $H = \mathrm{GL}(1|1) \times \mathrm{GL}(1|1)$, the coherent states do not transform as functions on G/H, but rather as sections of an associated line bundle, see Ref. 13 and references therein. In other words,

$$T_g |Z\rangle \neq |g \cdot Z\rangle.$$

The correct transformation law is derived as follows. We write the coherent states as $|Z\rangle = T_{s(Z,\widetilde{Z})}|0\rangle$, where

$$s(Z,\widetilde{Z}) = \begin{pmatrix} (1-Z\widetilde{Z})^{-1/2} & Z(1-\widetilde{Z}Z)^{-1/2} \\ \widetilde{Z}(1-Z\widetilde{Z})^{-1/2} & (1-\widetilde{Z}Z)^{-1/2} \end{pmatrix} = \begin{pmatrix} 1 & Z \\ 0 & 1 \end{pmatrix} \begin{pmatrix} (1-Z\widetilde{Z})^{+1/2} & 0 \\ 0 & (1-\widetilde{Z}Z)^{-1/2} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \widetilde{Z} & 1 \end{pmatrix}.$$

We then define an H-valued function $h(g;Z,\widetilde{Z})$ by

$$g s(Z,\widetilde{Z}) = s(g \cdot Z, g \cdot \widetilde{Z}) h(g;Z,\widetilde{Z}).$$

The explicit form of $h(g;Z,\widetilde{Z})$ can be found in Appendix B of Ref. 13. From (10) and the definition $T_g = \exp(\overline{c}_X(\ln g)_{XY}c_Y)$, one easily sees that the vacuum carries a one-dimensional representation μ of H:

$$T_h|0\rangle = |0\rangle\mu(h) \quad (\text{for } h \in H), \quad \mu(\mathrm{diag}(A,D)) = \mathrm{SDet}\, D^{-1}.$$

Therefore, the coherent states transform as

$$T_g|Z\rangle = T_g T_{s(Z,\widetilde{Z})}|0\rangle = T_{s(g\cdot Z, g\cdot \widetilde{Z})} T_{h(g;Z,\widetilde{Z})}|0\rangle = |g \cdot Z\rangle \mu(h(g;Z,\widetilde{Z})).$$

As a result, the R-matrix obeys the following transformation law:

$$\mathcal{R}(g \cdot \widetilde{Z}(o1), g \cdot \widetilde{Z}(o2), g \cdot Z(i1), g \cdot Z(i2))$$

$$= \mathcal{R}(\widetilde{Z}(o1), \widetilde{Z}(o2), Z(i1), Z(i2)) \prod_{\lambda=1,2} \mu(h(g;Z(o\lambda), \widetilde{Z}(o\lambda)))$$

$$\times \prod_{\nu=1,2} \mu(h(g;Z(i\nu), \widetilde{Z}(i\nu)))^{-1}.$$

Because each link is incoming with respect to one node, and outgoing with respect to another, the multipliers $\mu(h(g;Z,\widetilde{Z}))$ cancel when all R-matrices are multiplied together, so that the global $GL(2|2)$ invariance is recovered.

The transformation law for the R-matrix explains why a direct gradient expansion to extract $\sigma_{xy}^{(0)}$ from S_{latt} is difficult. Such an expansion locally produces terms such as

$$\mathrm{STr}(1-\widetilde{Z}Z)^{-1}\widetilde{Z}\partial_x Z(1-\widetilde{Z}Z)^{-1}\widetilde{Z}\partial_y Z,$$

for example, which are *singular* at $Z_{\mathrm{FF}}=\infty$ (the south pole on $M_{\mathrm{F}}=S^2$). When all of these terms are correctly summed over the entire network, they cancel, as is guaranteed by the global $GL(2|2)$ invariance, which permits us to rotate the south pole into any other point on the two-sphere S^2. However, the cancellation really does take place only *after* summation of terms. By the multiplier-corrected transformation law of the R-matrix, singular terms remain locally, making the extraction of the topological coupling difficult. This, then, is the reason why a gradient expansion was not attempted in Sec. VII B.

IX. ANISOTROPIC LIMIT

We have presented an analytical method for dealing with the Chalker–Coddington model in its original *isotropic* formulation, by mapping it on a lattice equivalent of Pruisken's nonlinear σ model. In this section we will review another way of arriving at Pruisken's model, a replica version of which was first published by D. H. Lee. Following Ref. 4 we now take for our starting point the *anisotropic* limit of the Chalker–Coddington model, and replace the unitary operator $U=U_1U_0$ by the Hamiltonian H for an array of chiral modes $n=1,2,...$, with velocity v and an alternating direction of propagation:

$$H=\oint dx\sum_{n,n'}\Psi_n^{\dagger}(x)[\delta_{nn'}(-1)^n iv\,\partial_x+V_{nn'}(x)]\Psi_{n'}(x).$$

The functions $V_{nn'}(x)=\overline{V}_{n'n}(x)$ are uncorrelated Gaussian random variables with zero mean and variance,

$$\langle V_{nn'}(x)\overline{V}_{nn'}(x')\rangle=2(u_0\delta_{nn'}+u_1\delta_{n,n'+1}+u_1\delta_{n,n'-1})\delta(x-x').$$

The symbol $\oint$ means that we are using periodic boundary conditions in x.

To prepare the treatment of the general case, we shall first consider the case of a single chiral mode n. The supersymmetric generating functional for the retarded and advanced Green's functions of H is set up in the usual way, see Sec. IV. Ensemble averaging over the random potential $V(x)=V_{nn}(x)$ leads to the functional integral

$$\mathscr{Z}=\int\mathscr{D}(\psi,\overline{\psi})\exp\oint dx(\overline{\psi}(\Lambda v\,\partial_x-\varepsilon)\psi-u_0(\overline{\psi}\Lambda\psi)^2),$$

where ε is a positive infinitesimal. As before, $\psi_X(x)$ is a super ''spinor'' field with four components $X=+B$ (retarded Boson), $X=+F$ (retarded Fermion), $X=-B$ (advanced Boson), and $X=-F$ (advanced Fermion). The notation means $\overline{\psi}\Lambda\psi=\overline{\psi}_{+\sigma}\psi_{+\sigma}-\overline{\psi}_{-\tau}\psi_{-\tau}$. If the energy in the retarded and advanced sectors is different, $\omega=E_+-E_-\neq 0$, we need to add a term $i\omega\overline{\psi}\psi$ to the Lagrangian. To probe this field theory more generally, we may couple it to an external non-Abelian gauge field $A(x)\in\mathrm{Lie}(GL(2|2))$ and consider

$$\mathscr{Z}_{\mathrm{Dirac}}^u[A]:=\int\mathscr{D}(\psi,\overline{\psi})\exp\oint dx(\overline{\psi}\Lambda v(\partial_x+A)\psi-u(\overline{\psi}\Lambda\psi)^2).$$

We have temporarily set $u=u_0$ for notational simplicity. The special coupling to frequency is retrieved by putting $A=(i\omega-\varepsilon)\Lambda/v$, independent of x. Note that the generating functional $\mathscr{Z}_{\text{Dirac}}^u[A]$ is invariant under local gauge transformations,

$$\mathscr{Z}_{\text{Dirac}}^u[A]=\mathscr{Z}_{\text{Dirac}}^u[{}^hA], \quad {}^hA=hAh^{-1}+h\partial_x h^{-1},$$

where $h(x)\in\text{GL}(2|2)$ acts on the spinor field by $\psi\mapsto h\psi$ and $\bar{\psi}\Lambda\mapsto\bar{\psi}\Lambda h^{-1}$.

It turns out that one can write down another $0+1$-dimensional field theory that has the very same local gauge invariance. The field of this theory is the supermatrix $Q=g\Lambda g^{-1}$, which was defined in Sec. II and transforms as $Q\mapsto hQh^{-1}$. The generating functional is

$$\mathscr{Z}_{\text{WZ}}[A]:=\int\mathscr{D}Q\exp\oint dx\,\text{STr}\,\frac{\Lambda}{2}g^{-1}(\partial_x+A)g.$$

Clearly, this satisfies $\mathscr{Z}_{\text{WZ}}[A]=\mathscr{Z}_{\text{WZ}}[{}^hA]$. The linear derivative term in the action is of the Wess–Zumino type, i.e., it cannot be expressed in a globally nonsingular way in terms of Q, and is often called a Berry phase. The theory is well-defined because the ambiguity under right translations $g\mapsto gh_{\text{R}}$ $[h_{\text{R}}\Lambda h_{\text{R}}^{-1}=\Lambda$ or, equivalently, $h_{\text{R}}(x)\in\text{GL}(1|1)\times\text{GL}(1|1)]$ gives rise to a factor

$$\exp\,\oint dx\,\text{STr}\,\frac{\Lambda}{2}h_{\text{R}}^{-1}\partial_x h_{\text{R}}=\exp\,\oint dx\partial_x\,\text{STr}\,\frac{\Lambda}{2}\,\ln h_{\text{R}}=\exp 2\pi im=1,$$

which is unobservable in the functional integral. (For more details see Sec. 3.4 of Ref. 13.)

The local gauge invariance shared by $\mathscr{Z}_{\text{Dirac}}^u[A]$ and $\mathscr{Z}_{\text{WZ}}[A]$ suggests the existence of some relation between these theories. In fact, the following statement is true:

$$\lim_{u\to\infty}\mathscr{Z}_{\text{Dirac}}^u[A]=\mathscr{Z}_{\text{WZ}}[A]. \tag{12}$$

This identity can be viewed as a $0+1$-dimensional analog of non-Abelian bosonization in $1+1$ dimensions and, since A couples to $v\psi\bar{\psi}\Lambda$ and $g\Lambda g^{-1}/2=Q/2$ in the respective cases, amounts to the ''bosonization rule,''

$$v\psi\bar{\psi}\Lambda\xrightarrow{u\to\infty} Q/2.$$

We now briefly sketch the proof of the non-Abelian bosonization formula (12). By the local gauge invariance of both theories, it is sufficient to prove the equality for an x-independent gauge field A. Moreover, A may be taken to be a diagonal matrix. In this special case, it is easy to apply a method that was described at length in Ref. 22 and works as follows. As a first step, one identifies $\mathscr{Z}_{\text{Dirac}}^u[A]$ as the coherent state path integral of a supersymmetric Hubbard Hamiltonian for bosons and fermions. Then, one takes advantage of the limit $u\to\infty$, which enforces a Hubbard constraint reducing the (low energy) degrees of freedom to that of a single superspin. And finally, one sets up the coherent state path integral for the superspin Hamiltonian. The latter path integral turns out to be $\mathscr{Z}_{\text{WZ}}[A]$, which completes the proof.

This proof, although straightforward, has the disadvantage of being somewhat indirect. A more direct procedure is to decouple the interaction term $(\bar{\psi}\Lambda\psi)^2$ by introducing a Hubbard–Stratonovitch field Q coupling to $\psi\bar{\psi}\Lambda$ and then to integrate out ψ and $\bar{\psi}$. The effective action for Q is

$$S[Q]=\oint dx\,\text{STr}\left(-\frac{u}{4v^2}Q^2+\ln\left(\partial_x+A+\frac{uQ}{v^2}\right)\right).$$

The next step is to simplify the Q field functional integral by means of the saddle-point approximation, as a result of which Q gets restricted to the nonlinear space $Q = g\Lambda g^{-1}$. This step, while only approximate in general, here becomes *exact* in the limit $u \to \infty$. [What makes this possible is the *stationarity* of the average density of states of $vi\partial_x + V(x)$.] By expanding $\ln(\Lambda g^{-1}(\partial_x + A)g + u/v^2)$ to linear order in $g^{-1}(\partial_x + A)g$, one obtains the action of the Wess–Zumino functional $\mathscr{Z}_{\mathrm{WZ}}[A]$. Higher orders are suppressed by powers of $v^2/(L_x u)$, with L_x the system size.

Let us finally return to the case of many counterpropagating chiral modes that are coupled by hopping matrix elements between neighboring modes, with variance u_1. The Gaussian random hopping gives rise to an additional term in the Lagrangian,

$$L \mapsto L + 2u_1 \sum_n \ (\overline{\psi}_n \Lambda \psi_{n+1})(\overline{\psi}_{n+1} \Lambda \psi_n) = L + 2u_1 \sum_n \ \mathrm{STr}(\psi_n \overline{\psi}_n \Lambda)(\psi_{n+1} \overline{\psi}_{n+1} \Lambda).$$

By the bosonization rule $v\psi\overline{\psi}\Lambda \to Q/2$ for $u_0 \to \infty$, the additional term turns into $(u_1/2v^2)\Sigma_n \mathrm{STr}(Q_n Q_{n+1})$. The condition of validity of this step is $u_0 \gg u_1$. As a result we obtain the Q field action,

$$S[Q] = \oint dx \sum_n \ \mathrm{STr}\left((-1)^n \frac{\Lambda}{2} g_n^{-1} \partial_x g_n + \frac{u_1}{2v^2} Q_n Q_{n+1} \right).$$

By a standard calculation[23,22] this is the action of the coherent state path integral for a quantum superspin Hamiltonian,

$$H_{\mathrm{spin}} = \frac{2u_1}{v^2} \sum_n \ \sum_{XY} \ (-1)^{|Y|+1} S_n^{XY} S_{n+1}^{YX}, \tag{13}$$

where $S^{XY} = \overline{c}_X c_Y$, and the graded commutation relations of the Fock operators $c, \overline{c}$ were given in Sec. VIII. To reproduce the alternating sign of the Wess–Zumino term, we must alternate the definition of the Fock vacuum. On even sites ($n \in 2\mathbb{Z}$) the relations (10) apply, whereas on odd sites ($n \in 2\mathbb{Z}+1$) we have

$$\overline{c}_{+\mathrm{B}}|0\rangle = \overline{c}_{+\mathrm{F}}|0\rangle = c_{-\mathrm{B}}|0\rangle = c_{-\mathrm{F}}|0\rangle = 0, \tag{14}$$

instead. The alternating vacuum plays the same role as the Néel state for ordinary spin systems and means that the superspin chain is ''antiferromagnetic'' in character.

The Hamiltonian H_{spin} with translational invariant coupling $J = 2u_1/v^2 > 0$ was shown in Ref. 13 to represent the low energy limit of the quantum Hamiltonian of Pruisken's nonlinear σ model at criticality. Thus, we finally conclude that the anisotropic Chalker–Coddington model with homogeneous (on average) inter-channel hopping is in the same universality class as Pruisken's model at $\sigma_{xy}^{(0)} = 1/2$. (Clearly, this line of reasoning is much less direct than that presented for the isotropic model in Secs. IV–VII.)

To conclude this section let me mention that the non-Abelian bosonization formula (12) extends to N channels:

$$\int \mathscr{D}Q \ \exp \ \oint dx \ N\mathrm{STr} \frac{\Lambda}{2} g^{-1}(\partial_x + A)g$$

$$= \lim_{u \to \infty} \int \mathscr{D}(\psi, \overline{\psi}) \exp \ \oint dx (\overline{\psi}_n \Lambda (\partial_x + A)\psi_n - u(\overline{\psi}_n \Lambda \psi_{n'})(\overline{\psi}_{n'} \Lambda \psi_n)).$$

The strategy of the proof is the same as for $N=1$. Using this formula we can easily reproduce the N-channel result $\sigma_{xy}^{(0)}=N/2$ of Sec. VII C. The N-channel bosonization formula offers also a quick way of analysing the zero-dimensional limit of the $2d$ chiral metal.[22]

X. MODIFIED NETWORK MODEL

The primary goal of all field theoretic analysis of the plateau-to-plateau transition in integer quantum Hall systems must be to *identify the fixed point theory* that describes this transition and uncover the *conformal structure* it is expected to have. Recent attempts[24] in this direction started from the observation that Pruisken's model or, rather, the closely related Dirac theory with random mass, random scalar potential and random gauge field, has a global $GL(2|2)$ symmetry for the case of one retarded and one advanced Green's function. [A finite imaginary part of the energy argument of the Green's functions reduces this symmetry to $GL(1|1)\times GL(1|1)$. However, in Secs. III and IV it is shown how such a symmetry breaking can be avoided by calculating a conductance between interior contacts of the network model.] This symmetry was then assumed to be promoted in the infrared to a Kac–Moody symmetry, which severely restricts the number of possible candidates for the fixed point theory. Unfortunately, these attempts have not been successful so far. What is needed as additional input to such considerations, which are purely algebraic, is a firm understanding of the Hilbert space structure, or the representations involved. It is one of the aims of the present work to contribute to such an understanding.

Following up on unpublished work by Read, it was argued in Ref. 13 that the quantum Hamiltonian of the critical theory should be a superspin Hamiltonian of the type (13) acting on a space of states built from alternating $GL(2|2)$ modules,

$$\ldots \otimes V\otimes V^*\otimes V\otimes V^*\otimes \ldots \, ,$$

where V and V^* are generated by the action $g\mapsto T_g=\exp(\overline{c}_X(\ln g)_{XY}c_Y)$ of $GL(2|2)$ on the vacua (10) and (14), respectively. (As follows from Sec. 4.4 of Ref. 13, the elements of V and V^* have an interpretation as holomorphic and antiholomorphic sections of a line bundle associated to $G\rightarrow G/H$ by the one-dimensional representation μ of H. This permits the construction of a nondegenerate pairing between V and V^*, so that these spaces can be viewed as being dual to each other, as suggested by our notation.) It is then natural to ask whether one might be able to construct an *integrable* superspin Hamiltonian, offering the possibility of an analytical and exact computation of critical properties. This question will now be addressed in the light of the results of Secs. VIII and IX.

Recall the ''functional'' vertex model presentation (7), (8) of the Chalker–Coddington model,

$$\mathscr{Z}=\int \mathscr{D}(Z,\widetilde{Z})\prod_{\text{nodes } n} R(\widetilde{Z}(u_n),\widetilde{Z}(d_n),Z(l_n),Z(r_n)),$$

$$R(\widetilde{Z}(1),\widetilde{Z}(2),Z(3),Z(4))=\langle Z(1)\otimes Z(2)|\hat{R}|Z(3)\otimes Z(4)\rangle,$$

$$\hat{R}=\exp(\overline{c}_X(\mu)(\ln \hat{U}_1)(\mu,\nu)c_X(\nu)). \tag{15}$$

(Note the change in the notation for links from capital to small letters.) It is instructive to pass from the *integration* over fields $Z,\widetilde{Z}$ to a *summation* over discrete degrees of freedom, as follows. Every link l emanates from one node, and ends at one node. Therefore, each $Z(l)$ occurs once in the ''bra'' and once in the ''ket'' of some R-matrix. Recall from Sec. VIII that P_V denotes the projector from Bose–Fermi Fock space onto the $GL(2|2)$ module V, where the elements of V, referred to as superspin states, are labeled by three quantum numbers $n_{f\pm}\in\{0,1\}$ and $n_{b+}=0,1,2...,\infty$. On using the closure relation,
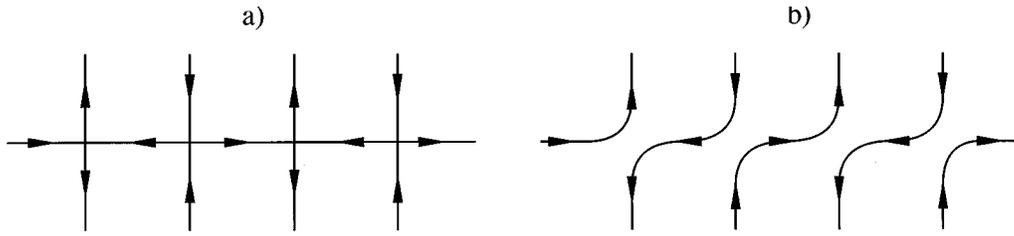
FIG. 4. (a) Illustration of the row-to-row transfer matrix of the supersymmetric vertex model associated with the one-channel Chalker–Coddington model. (b) The transfer matrix of the anisotropic model at $p=0$ translates the system by one lattice unit.

$$\int \prod_{\text{links } l} D(Z(l), \widetilde{Z}(l)) \otimes_l |Z(l)\rangle\langle Z(l)| = \otimes_l P_V(l),$$

the lattice functional integral over $Z, \widetilde{Z}$ turns into a partition sum over superspin configurations $\{\alpha(l)\}$ [with $\alpha(l) \in V(l)$] of a vertex model defined over the tensor product space $\otimes_l V(l)$:

$$\mathscr{Z} = \sum_{\{\alpha(l)\}} (-1)^{N_F} \prod_{\text{nodes } n} R_{\alpha(u_n)\alpha(d_n), \alpha(l_n)\alpha(r_n)}, \quad R_{\alpha\beta, \gamma\delta} = \langle \alpha \otimes \beta | \hat{R} | \gamma \otimes \delta \rangle,$$

where $(-1)^{N_F}$ is a sign factor due to supersymmetry (the partition sum is a supertrace), and $\hat{R}: V(l) \otimes V(r) \to V(u) \otimes V(d)$ is still given by (15). For $p=0$ (left turns only), $\hat{R}$ can be seen to be the identity map, while for $p=1$ (right turns only) we have $\hat{R} = \mathscr{P}$, where

$$\mathscr{P} | \alpha \otimes \beta \rangle = (-1)^{|\alpha||\beta|} | \beta \otimes \alpha \rangle$$

is the graded permutation operator.

Let me mention in passing that the presentation as a superspin partition sum can also be obtained from the network model *directly*,[25] without passing through the intermediate stage of a Z field formulation.

How is the model built from the vertex $R_{\alpha\beta, \gamma\delta}$ related to the superspin Hamiltonian (13)? Given $R_{\alpha\beta, \gamma\delta}$ we can set up the row-to-row transfer matrix, T, illustrated in Fig. 4(a), by summing over the superspin degrees of freedom that are situated on the horizontal links. (The meaning of the arrows in the present context will become clear below.) The corresponding Hamiltonian, defined as the logarithm of the transfer matrix, is nonlocal in general. A local Hamiltonian results on making the following modification of the isotropic network model. On nodes with coordinates $(n_x, n_y) \in \mathbb{Z}^2$ such that $n_x + n_y \in 2\mathbb{Z}$, the R-matrix is taken as it stands; but on the other half of the nodes $(n_x + n_y \in 2\mathbb{Z} + 1)$, we replace the left-right asymmetry parameter p by its complement $1 - p$. (At the level of the random network model, this is precisely what one does to arrive at the anisotropic limit of Sec. IX.) The row-to-row transfer matrix of the resulting *anisotropic* vertex model has the property, for $p=0$, of translating the system by one lattice unit [Fig. 4(b)]. By differentiating the logarithm of this transfer matrix at $p=0$, one gets a superspin Hamiltonian that couples only nearest neighbors and is precisely the Hamiltonian H_{spin} of (13).

H_{spin} is not expected to be exactly solvable. However, one may ask whether it could be made so by slightly changing some parameters while keeping the general structure and symmetries the same. To get some hint, we turn to the well-developed theory of integrable systems.[26] There, the integrability of a 1d quantum Hamiltonian H is traced back to the existence of a transfer matrix $T(u)$ that depends on a ''spectral'' parameter u in such a way that $T(u)T(v) = T(v)T(u)$ for all
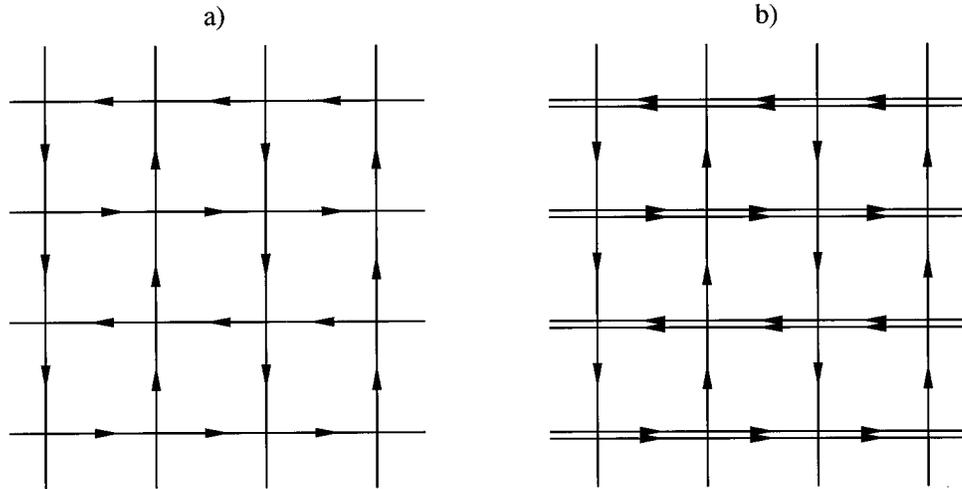
FIG. 5. (a) One-channel Chalker–Coddington model, modified so as to make the direction of motion invariant along every horizontal and vertical line. (b) The model of (a) modified further, by doubling the number of channels on the horizontal links.

u,v, and H is the logarithmic derivative of $T(u)$ at some special point $u=u_0$. A sufficient condition for $T(u)$ to form a commuting family is known to be the quantum Yang–Baxter equation for the R-matrix.

To my knowledge, most (if not all) of the integrable models discussed in the literature have one characteristic feature in common: their commuting family of transfer matrices possesses a ''classical'' limit, where T becomes the identity. In contrast, for the row-to-row transfer matrix associated to the Chalker–Coddington model [Fig. 4(a)], no such limit exists. The reason is simply this. The R-matrix $\hat{R}:V\otimes V\to V\otimes V$ in (8) relates incoming channels on horizontal (or vertical) links to outgoing channels on vertical (or horizontal) links. On the other hand, the transfer matrix T propagates the degrees of freedom from one row to the next. Therefore, to construct the row-to-row transfer matrix from the R-matrix, we must reinterpret some initial states as final states, and vice versa. This is done by noting that the space of linear maps $V\to V$ is isomorphic to the tensor product $V\otimes V^*$. In this way, one sees that in Fig. 4(a) vertical links with an arrow pointing up carry the space V, whereas vertical links with an arrow pointing down carry the dual space V^*. (This, then, is the meaning of the arrows in that figure.) Hence, the row-to-row transfer matrix of the SUSY reformulated Chalker–Coddington model is a map

$$T:\cdots\otimes V\otimes V^*\otimes V\otimes\cdots\to\cdots\otimes V^*\otimes V\otimes V^*\otimes\cdots,$$

which connects *inequivalent* spaces. Therefore, T cannot ever be the identity. This means that the Chalker–Coddington model lies outside the category of vertex models for which the well-developed Yang–Baxter machinery applies. Thus it seems that there exists no known systematic way of deforming the Chalker–Coddington model to integrability and obtain an analytical solution.

The above discussion, though being a falsification, also suggests a remedy. Given that the standard formalism of the theory of integrable systems requires the row-to-row transfer matrix to be a map,

$$T:\cdots\otimes V\otimes V^*\otimes V\otimes\cdots\to\cdots\otimes V\otimes V^*\otimes V\otimes\cdots, \tag{16}$$

we can turn things around and modify the network model accordingly. Consider the network shown in Fig. 5(a). As before, the electron follows the direction of motion dictated by the arrows. It picks up a random U(1) phase while propagating along the links, and is scattered deterministically at the nodes, just as for the original one-channel Chalker–Coddington model. The crucial difference from Fig. 2 is that now the direction of motion *does not alternate but is invariant* along vertical and horizontal lines. An electron incident on a node either passes straight through it, with probability p, say, or else is scattered to the right or left, as the case may be, with probability $1-p$. (Note that such a modification of the scattering dynamics has no justification from a microscopic picture of guiding center drift along equipotentials. However, since our aim is only to describe the *critical* behavior, we should have a certain amount of freedom in the choice of model.)

By construction, the row-to-row transfer matrix of the supersymmetric vertex model associated with the network model of Fig. 5(a) is a map of the desired type (16). One may now hope to be able to deform and extend this transfer matrix to a one-parameter family of commuting transfer matrices. In somewhat more detail, this hope is based on the following facts. Consider the Lie superalgebra $\mathscr{G}$ of polynomial maps $u \in \mathbb{C} \to \mathrm{gl}(2,2)$ (with u being the spectral parameter). If E_{AB} are the canonical generators of $\mathrm{gl}(2,2)$, the classical r-matrix $r(u,v) = (u-v)^{-1}\sum_{AB}E_{AB}(-1)^{|B|}E_{BA}$ gives rise to a co(super)commutator $\mathscr{G} \to \mathscr{G} \otimes \mathscr{G}$ in the usual way,[16] thereby turning $\mathscr{G}$ into a Lie bisuperalgebra. Quantization of $\mathscr{G}$ leads to a *Yangian superalgebra* $\mathscr{Y} := Y(\mathrm{gl}(2,2))$, which is a $\mathbb{Z}_2$-graded Hopf algebra deformation of the universal enveloping algebra of $\mathrm{gl}(2,2)$. Let $\Delta: \mathscr{Y} \to \mathscr{Y} \otimes \mathscr{Y}$ be the comultiplication of $\mathscr{Y}$, and $\Delta^{\mathrm{op}} = \mathscr{P} \circ \Delta$ its opposite. According to general principles,[27] there exists a formal object called the universal R-matrix $\mathscr{R} \in \mathscr{Y} \otimes \mathscr{Y}$, which is determined by the intertwining relation $\Delta^{\mathrm{op}}(a) = \mathscr{R}\Delta(a)\mathscr{R}^{-1}$ and has the expansion

$$\ln \mathscr{R}(u) = u^{-1}\sum_{AB}(-1)^{|B|}E_{AB} \otimes E_{BA} + \mathscr{O}(u^{-3}).$$

An irreducible matrix representation ρ of $\mathscr{Y}$ yields an R-matrix $R_\rho(u) = (\rho \otimes \rho)(\mathscr{R}(u))$ that is a rational function of u and solves the quantum Yang–Baxter equation.

When applying this formalism to our problem, we should beware of potential problems due to the *infinite-dimensionality* of the spaces V and V^*. Nevertheless, it does not seem unreasonable to expect the existence of R-matrices

$$R(u) = (\rho_V \otimes \rho_V)(\mathscr{R}(u)): V \otimes V \to V \otimes V,$$

$$\widetilde{R}(u) = (\rho_V \otimes \rho_{V*})(\mathscr{R}(u)): V \otimes V^* \to V \otimes V^*.$$

I have been able to verify that this expectation is fulfilled in the case of $R(u)$, by explicit construction. The existence of $\widetilde{R}(u)$ remains an open question at the present time. (Help from experts on quantum groups would be very much appreciated.) The latter case is more complicated to treat because the decomposition of the tensor product $V \otimes V^*$ into $\mathrm{gl}(2,2)$-irreducible subspaces involves a continuous series of representations of $\mathrm{gl}(2,2)$ (see Sec. 5.2 of Ref. 13), while in $V \otimes V$ only a discrete series appears.

Suppose now that both $R(u)$ and $\widetilde{R}(u)$ exist, at least in some domain of the spectral parameter u. Then, since the quantum Yang–Baxter equation is automatically fulfilled, we can build a one-parameter family of commuting transfer matrices. The next question is: given the modifications we have made, is the physics of such a model still that of the plateau transition? Surely, obtaining an exact solution remains a far goal. It is therefore helpful that the mapping onto Pruisken's nonlinear σ model provides us with a quick way to get oriented in the enlarged landscape of modified network models. Consider the model of Fig. 5(a) with parameter $p = 1/2$. Taking the continuum limit and computing the topological coupling in the same way as in Sec.

VII B, one finds $\sigma_{xy}^{(0)} = 0$ (mod 1). This means that the model does *not* lie in the quantum Hall universality class, but is in a massive phase with a finite (albeit large) localization length and exponentially decaying correlations.

This result could have been anticipated from the following heuristic argument. Imagine separating the network into two independent subsystems A and B, one consisting of the vertical lines and the other of the horizontal ones. Then couple the modes within each subsystem by weak tunneling amplitudes. What you get in this way are two copies of Lee's anisotropic limit of the Chalker–Coddington model. By the reasoning reviewed in Sec. IX, each of these is critical, with the σ model topological coupling constants being $\sigma_{xy,A}^{(0)} = \sigma_{xy,B}^{(0)} = 1/2$. Now join the two subsystems to form the network model of Fig. 5(a). From the meaning of L_{top} as a topological *density*, it is reasonable to expect that $\sigma_{xy}^{(0)} = \sigma_{xy,A}^{(0)} + \sigma_{xy,B}^{(0)} = 2 \times 1/2 = 0$ (mod 1) for the coupled system, if the fusion is done in such a way that $\sigma_{xx}^{(0)}$ for the coupled system is spatially homogeneous. Thus the network model of Fig. 5(a) will be noncritical at $p = 1/2$. (Its correlation functions at large scales should be similar to those of the two-channel Chalker–Coddington model at the symmetric point.) Since noncriticality is a generic property, this will remain so in a neighborhood of the point $p = 1/2$.

In view of this heuristic argument, we expect that criticality can be restored by superimposing on the noncritical network of Fig. 5(a) yet *another* copy of the anisotropic network model, thereby producing the network of Fig. 5(b). There, the horizontal lines carry two, rather than one, channels per link. The propagation on vertical links is governed by random U(1) phases, as before, but the horizontal links now carry random U(2) matrices (just like the two-channel Chalker-Coddington model). One of the horizontal channels passes straight through the nodes, the other one is subject to the rules specified for the model of Fig. 5(a). The topological coupling of the σ model is then found to have the critical value $\sigma_{xy}^{(0)} = 3 \times 1/2 (\text{mod } 1) = 1/2$. Thus the last, doubly modified network model is critical and, on symmetry grounds, lies in the quantum Hall universality class. From what was said above, it is also a suitable starting point for attempting to deform to an integrable model.

XI. SUMMARY

Several messages result from the present paper. First of all, a close relation between two standard models of the integer quantum Hall transition, namely the Chalker–Coddington model at its symmetric point $p = 1/2$, and the supersymmetric formulation of Pruisken's nonlinear σ model at $\theta = 2 \pi \sigma_{xy}^{(0)} = \pi$ (mod 2π), was established. Let us put this result in the proper context. To be sure, it has been clear for a number of years now that some sort of relation between these models ought to exist. We know that both are critical and belong to the same universality class, so they cannot but describe the same physics at long wave lengths. However, prior to our work, the understanding of the precise connection between Chalker–Coddington and Pruisken was rather indirect, relying on a double use of the anisotropic (or Hamiltonian) limit. The connection went as follows. At one end, by taking the Chalker–Coddington model and going to its anisotropic limit, D. H. Lee arrived at a network model consisting of an array of chiral modes with an alternating direction of motion. At the other end, the anisotropic limit of Pruisken's model was investigated, by an elaboration of the work of Shankar and Read[28] on the O(3) nonlinear σ model. It was argued in Ref. 13 that the Hamiltonian limit of Pruisken's model at $\theta = \pi$ and strong coupling (small $\sigma_{xx}^{(0)}$) is an antiferromagnetic superspin chain. Now, the array of chiral modes and the superspin chain are easy to relate by conventional techniques. At the level of the replica trick that this was done in Ref. 4, the correct supersymmetric extension follows from Ref. 22. In more detail, the functional integral representation of the array of chiral modes maps on a nonlinear σ model with an alternating sum of Wess–Zumino terms. (As we have seen, this mapping is based on a $0+1$- dimensional analog of non-Abelian bosonization in $1+1$ dimensions.) The latter, in turn, coincides with the coherent-state path integral of the antiferromagnetic superspin chain.

One disadvantage of the above way of relating the models is that one does not have good control over the numerical value of the coupling constant $\sigma_{xx}^{(0)}$. In the present work, this uncertainty was resolved by dealing directly with the *isotropic* models, using a novel scheme devised in Ref. 15. What we have shown is this. Starting from the Chalker–Coddington model at $p = 1/2$ and *doing no more than an exact transformation followed by a continuum limit*, we arrive at Pruisken's model with $\theta = \pi$ and $\sigma_{xx}^{(0)} = 1/4$. Or, in different words, the former model is equivalent to a specific lattice discretization of the latter.

Previously, Pruisken's model was thought to be associated primarily with the white noise limit $l_c \ll l_B$, which is where Pruisken's derivation begins. Recall, though, that in the course of deriving his model, Pruisken made a saddle-point approximation to eliminate the ''massive'' modes. Naive use of this approximation scheme is justified only in the limit of large $\sigma_{xx}^{(0)}$ (high Landau level). In contrast, the present work makes *no* such approximation. The only assumption we needed was the dominance of slowly varying fields in the functional integral, allowing us to pass from the lattice to the continuum. Thus, contrary to what might have been expected, Pruisken's model at small $\sigma_{xx}^{(0)}$ is actually associated more closely with the *high-field limit*, $l_c \gg l_B$, as it is this limit that provides the microscopic justification of the network model. Note, however, that the ratio of microscopic length scales l_c / l_B is expected to be an irrelevant parameter at a critical point with an infinite correlation length. Thus, our result is not in conflict with Pruisken's derivation of the σ model as a critical theory.

In my opinion, neither the Chalker–Coddington model nor Pruisken's model hold much promise for an exact analytical solution in the near future. If so, the mapping of one model on the other is not yet a big step forward. The good news is that there are several useful spinoffs. D. K. K. Lee and Chalker suggested modeling the random flux problem (i.e., the motion of a single electron in a random magnetic field) by a network with two channels per link and local U(2) gauge invariance. Our mapping onto a nonlinear σ model easily extends to include this case. The coupling constants of the continuum field theory were found to be $\sigma_{xx}^{(0)} = 2 \times 1/4 = 1/2$ and $\sigma_{xy}^{(0)} = 2 \times 1/2$ (mod 1) $= 0$ at the symmetric point of the random flux problem. This is a massive theory with exponential decay of all correlation functions are large scales. Thus, contrary to claims made in the literature, there is no room for truly extended states in the random flux problem, at least not by slight deformation away from the two-channel network model. This conclusion had already been reached in Refs. 21 and 4.

Another spinoff helps us along in our quest to understand the integer quantum Hall transition. We observed that the SUSY reformulated Chalker–Coddington network model has the structure of what is known as a vertex model in statistical physics. Motivated by information from the theory of integrable systems, we then modified the network model in several ways. First, we abandoned the alternating direction of the electron's motion along the horizontal and vertical straight lines of the network. Instead, we took the direction of motion to be constant along every such line. The mapping onto Pruisken's model indicates that this modification changes the physics: the value of the topological coupling now is $\sigma_{xy}^{(0)} = 0$, which corresponds to a noncritical state. We argued that criticality can be restored by doubling the number of channels on all horizontal links (or on all vertical links). Alternatively, we can spatially separate the two channels on horizontal links and return to a model with only one channel per link, at the expense of doubling the size of the unit cell in the vertical direction.

The resulting modified network model is critical, but no longer has a justification from a microscopic picture of guiding center drift along equipotentials. Its virtue is that the corresponding supersymmetric model is of a type for which systematic ways of constructing solutions of the quantum Yang–Baxter equation are in principle available. Whether our model can actually be deformed into one that is a Yang–Baxter integrable, is a question whose answer lies far beyond the scope of the present paper.

ACKNOWLEDGMENTS

Most of this work was done while the author was visiting the Institute for Theoretical Physics, UCSB, Santa Barbara, and the Aspen Center for Physics. Useful discussions with John Chalker, Andreas Ludwig, and Boris Muzykantskii are acknowledged.

[1] B. Huckestein, ''Scaling theory of the integer quantum Hall effect,'' Rev. Mod. Phys. 67, 357–396 (1995).

[2] S.A. Trugman, ''Localization, percolation, and the quantum Hall effect,'' Phys. Rev. B **27**, 7539–7546 (1983).

[3] J.T. Chalker and P.D. Coddington, ''Percolation, quantum tunneling, and the integer quantum Hall effect,'' J. Phys. C **21**, 2665–2679 (1988).

[4] D.H. Lee, ''Network models of quantum percolation and their field-theory representation,'' Phys. Rev. B **50**, 10788–10791 (1994).

[5] C.-M. Ho and J.T. Chalker, ''Models for the integer quantum Hall effect: the network model, the Dirac equation, and a tight-binding Hamiltonian,'' cond-mat/9605073.

[6] F. Wegner, ''The mobility edge problem: Continuous symmetry and a conjecture,'' Z. Phys. B **35**, 207–210 (1979).

[7] K.B. Efetov, ''Supersymmetry and theory of disordered metals,'' Adv. Phys. **32**, 53–127 (1983).

[8] A.M.M. Pruisken, ''On localization in the theory of the quantized Hall effect: A two-dimensional realization of the θ-vacuum,'' Nucl. Phys. B **235**, 277–298 (1984).

[9] A.M.M. Pruisken, ''Quasi particles in the theory of the integral quantum Hall effect (II). Renormalization of the Hall conductance or instanton angle theta,'' Nucl. Phys. B **290**, 61–86 (1987).

[10] H. Levine, S.B. Libby, and A.M.M. Pruisken, ''Theory of the quantized Hall effect (I),'' Nucl. Phys. B **240**, 30–40 (1984).

[11] H.A. Weidenmüller, ''Single electron in a random potential and a strong magnetic field,'' Nucl. Phys. B **290**, 87–110 (1987).

[12] D.E. Khmel'nitskii, ''Quantization of Hall conductivity,'' JETP Lett. **38**, 552–556 (1983).

[13] M.R. Zirnbauer, ''Towards a theory of the integer quantum Hall transition: From the nonlinear sigma model to superspin chains,'' Ann. Phys. (Leipzig) **3**, 513–577 (1994); cond-mat/9410040.

[14] A.M. Perelomov, *Generalized Coherent States and Their Applications* (Springer-Verlag, Berlin, 1986).

[15] M.R. Zirnbauer, ''Supersymmetry for systems with unitary disorder: Circular ensembles,'' J. Phys. A **29**, 7113–7136 (1996).

[16] V. Chari and A. Pressley, *A Guide to Quantum Groups* (Cambridge University Press, Cambridge, 1994).

[17] J.J.M. Verbaarschot and M.R. Zirnbauer, ''Critique of the replica trick,'' J. Phys. A **17**, 1093–1109 (1985).

[18] H.A. Weidenmüller and M.R. Zirnbauer, ''Instanton approximation to the graded nonlinear sigma model for the integer quantum Hall effect,'' Nucl. Phys. B **305**, 339–366 (1988).

[19] M.J. Rothstein, ''Integration on noncompact supermanifolds,'' Trans. Am. Math. Soc. **299**, 387–396 (1987).

[20] M.R. Zirnbauer, ''Riemannian symmetric superspaces and their origin in random matrix theory,'' J. Math. Phys. **37**, 4986–5018 (1996).

[21] D.K.K. Lee and J.T. Chalker, ''Unified model for two localization problems: Electron states in spin-degenerate Landau levels and in a random magnetic field,'' Phys. Rev. Lett. **72**, 1510–1513 (1994).

[22] L. Balents, M.P.A. Fisher, and M.R. Zirnbauer, ''Chiral metal as a ferromagnetic super spin chain,'' cond-mat/9608049.

[23] E. Fradkin, *Field Theories of Condensed Matter Systems* (Addison-Wesley, Redwood City, 1991).

[24] C. Mudry, C. Chamon, and X.G. Wen, ''Two-dimensional conformal field theory for disordered systems at criticality,'' Nucl. Phys. B **466**, 383–443 (1996).

[25] I.A. Gruzberg, N. Read, and S. Sachdev, ''Scaling and crossover functions for the conductance in the directed network model of edge states,'' cond-mat/9612038.

[26] R.J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Academic, London, 1982).

[27] V.G. Drinfel'd, ''Quantum groups,'' in *Proceedings of the I.C.M.,* Berkeley, 1986 (American Mathematical Society, Providence, RI, 1987), pp. 798–820.

[28] R. Shankar and N. Read, ''The $\theta = \pi$ nonlinear σ model is massless,'' Nucl. Phys. B **336**, 457–474 (1990).

Erratum: ''Toward a theory of the integer quantum Hall transition: Continuum limit of the Chalker–Coddington model'' [J. Math. Phys. 38, 2007 (1997)]

Martin R. Zirnbauer

Universitat zu Koln, Institut fur Theoretische Physik, Zulpicher Str. 77, 50937 Koln, Germany

The value of the coupling constant $\sigma_{xx}^{(0)}$ computed in Sec. VII A is not correct. Its calculation was based on the reasonable but false assumption that the field configurations which dominate at large wavelengths are smooth. As a matter of fact, the relevant configurations contain a small high-momentum component due to the alternating structure of the Chalker–Coddington network. When this is taken into account, the value of the longitudinal conductivity changes to $\sigma_{xx}^{(0)}=1/2$ for the single-channel model and $\sigma_{xx}^{(0)}=N/2$ for the N-channel model. The rest of the paper is unaffected by this change.