Delocalization, duality, and scaling in the quantum Hall system

C. A. Lütken* and G. G. Ross
Department of Physics, University Of Oxford, 1 Keble Road, Oxford, OX1 3NP, United Kingdom
(Received 11 September 1992)

We construct an effective-field theory for the quantum Hall system which embodies both localization and fractional statistics. The latter involves a Chern-Simons interaction, while the former involves a generalization of conventional localization theory. The theory is invariant under "complexified" duality transformations of the conductivities which appear as effective parameters of the model. By exploiting these parameter space symmetries, as well as the conformal symmetry which appears at renormalization-group fixed points, we are able to extract a precise prediction for the whole scaling diagram. It exhibits both fractional and integer phases, the exact location of all fixed points, and universal scaling exponents. With a plausible identification of the universality class of the theory in the replica limit, the value of the critical exponent for the delocalization transition between plateaus in the Hall conductivity is found to be $\frac{7}{3}$, in apparent agreement with available scaling experiments.

I. INTRODUCTION

A discussion of the quantum Hall system is conventionally split into three parts: the integer effect, the fractional effect, and the scaling properties of the transport coefficients in the transitions between the plateaus. Since this corresponds rather closely to the familiar progression of ideas from semiclassical to first-quantized and finally to second-quantized theory, this separation is valuable for developing physical intuition and also for obtaining some quantitative information about the plateaus, as shown most clearly in Laughlin's work. Nevertheless this trichotomy is artificial if one believes that it is possible to give a unified field-theoretical description of all macroscopic observables in the quantum Hall system. From this point of view it should be possible, and as far as the scaling properties go it is *essential,* to encode all of these aspects of charge transport in an external magnetic field in a more or less conventional effective-quantum-field theory, whose validity must ultimately be verified by making contact with the microscopic processes, especially localization, believed to be responsible for the quantum Hall effect.

Once this idea is entertained attention immediately shifts away from the striking quantization of the Hall conductivity, towards the critical phenomena signaled by the observed scaling in the transition regions between the plateaus. This is because the only way to accommodate scaling is in a second-quantized treatment, or its equivalent in statistical mechanics.

A remarkable development has been the experimental discovery that the transitions between both integer and fractional levels are described by the *same* critical exponent. This "universality of exponents," as opposed to ordinary universality which is concerned with the behavior near a single fixed point, must be a novel and fundamental property of the effective-field theory, and in a previous paper[1] we suggested that this phenomenon should be taken as evidence for a discrete symmetry acting on the parameter space of the theory. Such a symmetry can map all fixed points and associated scaling equations into each other, thus accounting for the observed "superuniversality" of the delocalization exponent. In view of the rather complicated "nested" or "hierarchical" structure of the fractional and integer levels, and thus the fixed points, it is fairly clear that such a symmetry group can be neither finite nor abelian. Surprisingly, a group of precisely the required type appears in a simple class of two-dimensional spin models consisting of two p-state Potts models ($p=2,3$) coupled through an antisymmetric term. It was shown in Ref. 1 how this symmetry relates all fixed points governing the critical behavior of all levels, both integer and fractional, and thus predicts both the level structure of the plateaus and a "superuniversal" scaling exponent for transitions between levels. The structure of the phase diagram and the properties of the renormalization-group (RG) flow of the system may, to a large extent, be determined by demanding consistency with this modular symmetry. Preliminary studies suggest that this structure is in good agreement with experiment.

While we initially[1] only offered this particular class of models as proof of the existence of dynamical systems with the desired symmetry properties, we shall in this paper argue that the effective-field theory describing the macroscopic observables of the quantum Hall system can indeed be mapped onto a coupled Potts model, suitably generalized by analytic continuation in p to $p=1$. These coupled Potts models encode the degrees of freedom which are relevant at the largest scales, i.e., the macroscopic degrees of freedom. The fact that we are using a lattice spin model to represent these may be confusing, since the microscopic degrees of freedom (the spins) of the spin model are not to be thought of as representing the true microscopic degrees of freedom, but should be regarded simply as a convenient way of encoding the correct *universality properties.* The macroscopic degrees of freedom are accompanied by effective parameters, in our case transport coefficients or response functions, which can be controlled by or related directly to experiment.

We must now try to show how this class of models emerges at large scales from the microphysics believed to be at work in these semiconductors, the two central concepts being localization and fractional statistics, and in the process demonstrate that the relevant member of the class has $p=1$. Since the diagonal or dissipative conductivity σ_{xx} vanishes on the plateaus found in the quantum Hall system, it is clear that, except at transition points, the leading term in the effective action should be associated with the transverse or Hall conductivity σ_{xy}. This is rather surprising since the standard theory of dissipative resistance, due to localization and impurity scattering, reveals that σ_{xx} parametrizes an ordinary kinetic term, which is usually considered to be the most important part of the effective action. However, in $2+1$ dimensions this is not true: the leading term is the Chern-Simons (CS) action. This is a topological invariant, i.e., it does not contribute to the equations of motion, and its physical meaning and significance have only recently been appreciated. Provided that the CS-gauge field (the *statistical* gauge field) is coupled to the charge carriers, the CS term changes the effective spin and statistics of these particles, leading to the possibility that the effective degrees of freedom (quasiparticles) in $2+1$ dimensions can be anyonic with fractional spin and charge. This possibility arises because the representations of the rotation group are continuous in two dimensions. That means[2,3] that the wave function of a system of identical particles cannot simply be labeled by their quantum numbers, but depends on the braiding characteristics of their path histories.

While it is possible to give a field-theoretic description of this directly in terms of the physical, microscopic degrees of freedom, the action is necessarily nonlocal and therefore extremely unpleasant. However, the path dependence of the quantum phase has the flavor of an Aharonov-Bohm effect, and it is a remarkable fact[4] that the action can be recast in a completely local form by introducing an auxiliary (fictitious) abelian gauge field, minimally coupled to the charged particles, with a CS self interaction. Just as the potential of a vanishing three-dimensional electromagnetic field carries information about the phase of the electron, the CS-potential miraculously manages to encode all the complicated phase information of a two-dimensional system of charge carriers.

In summary, we have three reasons for considering a CS term in our effective action: the microphysics admits fractional excitations, which are most conveniently encoded in a field theory by using the CS trick; it must appear at large scales if it can, because it has the lowest scaling dimension; and in a geometrical interpretation of quantum-field theory it appears *naturally* as the dominant term, because it encodes topological information. This may not be unrelated to it having the lowest scaling dimension, since one expects topological terms to contain the fewest derivatives ("only long wavelength modes are needed to explore the topology"), but it is not clear to us precisely what the connection is.

The field-theoretic representation of the dissipative conductivity σ_{xx} in a CS theory has recently been discussed by Kivelson, Lee, and Zhang.[5] Starting from an effective Landau-Ginzburg-type action describing the propagation properties of anyonic degrees of freedom, they integrated out the CS bosons and statistical gauge field to obtain expressions for the conductivity tensor parametrized by functions encoding the linear-response properties of the CS bosons. Assuming that these response functions are universal for the hierarchy of anyonic states that can be built on the original spin-polarized electron state comprising the full Landau level, they obtained relations between the conductivity tensors associated with different phases of the quantum Hall system. We shall show below that these relations follow from a subgroup of the full modular symmetry introduced in Ref. 1. Moreover, if the Chern-Simons response functions are independent of the anyonic spin, irrespective of whether they belong to the Halperin-Haldane hierarchy or not, the form of their conductivity tensor is such that the full modular symmetry is obtained.

While the analysis of Ref. 5 lends strong support to the idea[1] that modular symmetry, or a subgroup of it, is relevant to the description of the quantum Hall system, it does not provide a complete derivation of the symmetry, and sheds no light on the nature of the system at criticality, which presumably is governed by a conformal field theory. In addressing these questions below, because a theory of anyon localization does not yet exist we shall try to proceed cautiously, as close to the conventional theory of fermion localization as we can. However, in studying the propagation of charged anyonic states in the presence of impurities we find that a modification of the traditional treatment is necessary. The usual assumption is that the effect of impurities may be modeled by a pointlike scattering from a random potential with Gaussian distribution. This may be sufficient for sharply localized states, like electrons in ordinary conductors, but we do not believe that this is a good model when the charge carriers are nonlocal excitations, possibly with fractional charge and spin, arising from the cooperative physics in effectively two-dimensional semiconductors. Instead of having just the conventional interaction in the effective-field theory of localization, we shall argue that a more general form is needed, constrained only by the symmetries of the problem. The occurrence of a nonvanishing value for σ_{xx} corresponds to the appearance of delocalized states, which are related to states left massless by the effects of the impurity potential. In this the symmetry observed by Wegner,[6] relating the advanced and retarded propagators of the charge carriers, plays a vital role, for it identifies the relevant degrees of freedom which contain the massless states at criticality.

To summarize, the field-theoretical *bulk picture,* as extended below to include the localization of anyons, is apparently able to account for all available data. The traditional integer theory, both the first-quantized "Laughlin picture" of the plateaus and the "Wegner-Pruisken" field theory of (de)localization, is included, when the anyonic quasiparticles happen to be fermions. In general, however, the statistics of the quasiparticles in two dimensions may be fractional, in which case they are excited from the ground state of a fractional phase. The conventional hierarchy mechanism for generating fractions is also contained in this picture, being nothing else than the discrete symmetry restricted to the Hall regime. Laughlin's quasiparticle excitations are the anyons.

In this simple picture integer and fractional phases are treated on the same footing. Only one mechanism—localization—is needed to account for all data. The difference between the integer and fractional effects is "only" in the nature of the quasiparticle object being localized, not in the physical mechanism itself. While the degree of disorder dramatically effects the physical properties of the charge-carrying states, even the spin and statistics of the quasiparticles that can be excited in the system, the transport mechanism is always the same. So while the physics is different in different samples, it is always of the same *form,* which is why the scaling is always of the same form. *It is this form invariance which is encoded in the global discrete symmetry.*

Finally, we close this lengthy introduction by comparing and contrasting the "bulk picture" pursued in this paper, with the "edge picture." The question of the role of edge states is also related to localization, since they are always delocalized and therefore potentially charge carrying. The reason for the existence of plateaus is the same in either the bulk or the edge picture, being due to the quantized Landau levels and the energy gaps. There are two possible effects giving rise to the width of the Hall plateaus; the existence of edge states and the usual impurity broadening. The contribution of the pure edge states is inversely proportional to the width of the sample and is usually small. So in real samples the dependence of the plateau width on the impurity level will be almost the same as in the pure bulk description discussed in this paper. We note also that the quantum Hall effect persists even in the absence of edges as has recently been conclusively demonstrated in a Corbino disk experiment.[7]

We expect the delocalization transition between plateaus to be completely given by the bulk physics sensitive to impurities. The reason is that the edge states are nondissipative and thus cannot contribute to the appearance of a nonzero longitudinal conductance as one makes the transition between plateaus via delocalized states. The physics of these transitions must be determined by the bulk properties of the system which describe the dissipative states. Thus we expect the delocalization transition to be completely determined by the bulk physics sensitive to impurities. No purely one-dimensional "edge" mechanism has been suggested to account for this phenomenon.

In summary, we have attempted to construct the rudiments of a theory of localization of anyons, which we believe to be responsible for most or all of the observed properties of the quantum Hall effect. If either of the two central ideas encoded in this theory—localization (due to impurities) or exotic quasiparticle statistics (due to strong parity-violating interactions in a two-dimensional electron gas in a magnetic field)—turn out not to capture the essential physics in real samples, then our theory is not applicable, but so far there is no evidence to suggest this.

The outline of the paper is as follows. As an introduction to discrete parameter space symmetries ("complexified Kramers-Wannier duality"), and in order to fix notation and review results from conformal field theory which will be needed below (Sec. VI), in the next section we review the discrete *modular* symmetry of the coupled Potts models, and discuss its phase and renormalization-group structure.

Section III reviews the recent discussion of anyonic propagation,[5] and its relation to the discrete modular symmetry.

Section IV reviews the field-theoretic approach to delocalization of ordinary fermions and the Wegner symmetry[6] in a way which is appropriate to the discussion of localization in the quantum Hall system. This leads to an identification of the extended states with the coset space of fields associated with the Wegner symmetry.[8]

In Sec. V we discuss how this analysis is modified in the presence of Chern-Simons interactions, which must be included (generated) in an effective action in order to accommodate the anyonic states that can appear in two dimensions. These lead to the possibility of many more delocalization fixed points, with their associated extended states. We argue that the conventional choice of scattering potential is not good enough for discussing the localization of the collective nonlocal excitations that appear in the quantum Hall system. We show that the effect of allowing for a more general form of the impurity potential is to explicitly break the Wegner symmetry to a discrete $\mathbb{Z}_p$ subgroup. The associated Landau-Ginzburg theory with this symmetry is shown to lie in the universality class of the self-dual coupled Potts models, discussed in Sec. II. From this identification follows the modular symmetry relating the conductivity tensor associated with transitions between different levels, in agreement with the expectation (described in Sec. III) following from the analysis of Ref. 5. In this case, however, no assumptions are needed about the linear-response properties of the CS bosons, these being completely determined by the conformal field theory at the fixed point.

In Sec. VI we study the theory in the replica limit, and argue that this corresponds to a Potts model with $p \to 1$. We can then compute the delocalization exponents using standard results from conformal field theory, which show that this limit can be interpreted as a percolation problem. This is in good agreement with a microscopic picture of the Hall effect as due to "quantum percolation" of charge through the sample. The final step is to argue that the classical bond-percolation exponent $\frac{4}{3}$, obtained from the conformal field theory at the delocalization fixed points, is shifted by one due to tunneling effects between the geometrical percolation clusters. Thus, while the location of the RG fixed points are given by the $p=1$ self-dual theory, we expect the critical exponent to be $\frac{7}{3}$, in agreement with numerical work and recent experiments.

II. SCALING PROPERTIES OF SELF-DUAL MODELS

The "complexified Kramers-Wannier" symmetries, with which we will be concerned in the following, first appeared in an investigation of abelian lattice gauge theories which turned out to be intimately related to a particular type of coupled $\mathbb{Z}_p$-symmetric spin model.[9,10] In terms of two "fields" ϕ^a ($a=1,2$) restricted to take values in the integers mod p (representing $\mathbb{Z}_p$-valued spins $\mathbf{s}_i$ on dual lattices), the coupled spin model action is of the simplest possible form containing no more than two derivatives:

$$L_p = -\beta\delta_{\mu\nu}\delta_{ab}\partial^\mu\phi^a\partial^\nu\phi^b + i\alpha\epsilon_{\mu\nu}\epsilon_{ab}\partial^\mu\phi^a\partial^\nu\phi^b \quad (\beta>0)$$
$$= 2i(\tau\partial\varphi\bar{\partial}\bar{\varphi} - \bar{\tau}\partial\bar{\varphi}\bar{\partial}\varphi) . \qquad (1)$$

Here α and β are two real parameters, which in the second, complexified, form of L_p have been traded for one complex parameter $\tau=\alpha+i\beta$. The latter equality in (1) is obtained by introducing the complex coordinate $z=x+iy(\partial=\partial/\partial z)$, the complex scalar field $\varphi=\phi^1+i\phi^2$, and their complex conjugates (denoted by overbars).

In order to give meaning to this somewhat heuristic, but suggestive, "field-theoretic" representation of a coupled $\mathbb{Z}_p$-symmetric spin system, care must be exercised when interpreting the "lattice derivatives." If $(\partial\phi)^2$ means the difference in ϕ evaluated at neighboring points on the lattice ($\mathbf{s}_i\cdot\mathbf{s}_j$ in conventional spin language), then (1) is the coupled clock (also called planar or vector Potts) model, which was investigated most lucidly by Cardy and Rabinovici.[9,10] To compare (1) with their action, one must split the fields ϕ into the conventional "spin-wave" and "vortex" components. We have not done so here because we wish to emphasize the complex extension of the more familiar spin-wave-vortex duality. Alternatively, if $(\partial\phi)^2$ is nonvanishing only if neighboring field values coincide (δ_{s_i,s_j} in spin language), then (1) represents two p-state (standard) Potts models living on dual lattices, which communicate through the "spin-wave-vortex" coupling encoded in the antisymmetric term. If $p=2$ or 3 the clock and Potts models coincide exactly, but for other values of p they do not. In particular, while the clock model exhibits nontrivial critical behavior for all integer $p\geq 2$, but is not defined for other values of p, the Potts model has no second-order transition for $p>4$, but can be analytically continued to any real value $0\leq p\leq 4$. It is the universality class of the coupled ($p\rightarrow 1$)-state Potts model which appears to encode the large distance behavior of the quantum Hall system at criticality, i.e., at the delocalization transition where charge "percolates" through the macroscopic sample.

The neat appearance of τ in (1) suggests that the parameter space is the complex upper half plane $\mathbb{H}=\{\tau\in\mathbb{C}|\mathrm{Im}\tau>0\}$, and it is the symmetries of this space which will concern us in the following. As first shown by Cardy,[10] the partition function $\mathcal{Z}_p$ of (1) is invariant under modular transformations:

$$\mathcal{Z}_p(\gamma(\tau),\gamma(\bar{\tau})) = \mathcal{Z}_p(\tau,\tau)$$
$$= \mathrm{Tr}\exp\left\{-\int d^2z L_p(\tau,\bar{\tau})\right\} , \qquad (2)$$
$$\gamma\in\Gamma(1)\equiv \mathrm{SL}(2,\mathbb{Z}) .$$

The modular group $\Gamma(1)$ is generated by translations $T(\tau)=\tau+1$ and inversions $S(\tau)=-1/\tau$. The rich structure of this group is entirely due to the fact that S and T do not commute. Hence, if we consider only one spin model, so that no antisymmetric term is possible, then the only remnant of $\Gamma(1)$ is the real transformation $S(\beta)=1/\beta$, which is just the Kramers-Wannier transformation.

It is well known that this symmetry can be used to locate critical points of spin models.[11] $\mathcal{Z}_p(i\beta)$ reduces to a single $\mathbb{Z}_p$ spin model in the thermodynamic limit ($p=2$ is the Ising model). In the Potts case we know[12] that there is a unique second-order phase transition at the self-dual point ($\tau=i$, i.e., $\beta=1$) when $p\leq 4$, and that the transition is first order when $p>4$. This result was generalized in Refs. 9 and 10 to the coupled clock models. A simple comparison of energy and entropy showed that the phase-boundary extends in a unique way away from $\tau=i$, provided that $p<2\sqrt{3}$. The rest of the phase boundary is then uniquely fixed by modular invariance, so that for $p=2,3$ (coupled Ising and 3-state Potts models) the phase diagram is given by the tree of solid lines shown in Fig. 1. [The flow lines on the tree in Fig. 1 do not correspond to any of Cardy's models, but to a coupled ($p<2$)-state Potts model.] For higher values of p a new "Coulomb" phase injects itself along this phase boundary, and for large p it grows to dominate the phase diagram.

From the fact that the phases of the diagram in Fig. 1 only touch the real axis at fractional values, it is immediately clear why $\Gamma(1)$ is a promising group: if τ can be identified with the *complexified conductivity* $\sigma=\sigma_{xy}+i\sigma_{xx}$, then σ_{xy} will be forced by the phase diagram alone to take fractional values when σ_{xx} vanishes. Furthermore, as shown in Ref. 1, the location of the bifurcation points where new fractional phases become possible as σ_{xx} is reduced agree remarkably well with available scaling data.

The bifurcation points are obviously fixed points of the renormalization group since RG flows cannot cross phase boundaries. That they are also fixed points of $\Gamma(1)$ follows from the fact that the modular group is the free product of $\mathbb{Z}_2$, generated by S, and $\mathbb{Z}_3$, generated by TS. This implies that there are two types of fixed points on the tree, of order 2 ($S^2=1$) and 3 [$(TS)^3=1$]. The bifurcation point at $\tau=j\equiv\exp(\pi i/3)$ is an "elliptic" fixed point of order 3 ($\ominus$ in Fig. 1), because $TS(j)=(TS)^2(j)=j$. Since modular transformations are conformal, angles are preserved under TS and the phase boundaries meeting at j must do so at an angle of $2\pi/3$. For the same reason every image $\gamma(j)$ [$\gamma\in\Gamma(1)$] of j is also a triskelion, i.e., a fixed point of order 3. Similarly, $\tau=i$ is an "elliptic" fixed point of order 2 ($\otimes$ in Fig. 1), because $S(i)=i$, and so are all its images $\gamma(i)$. The latter are natural candidates for delocalization fixed points, and this also fits well with scaling data on the transition between many integer levels.

In addition, there are two other types of fixed points not located on the self-dual tree. Strictly speaking, they do not lie in the parameter space $\mathbb{H}$ at all, but on its compactification $\bar{\mathbb{H}}=\mathbb{H}\cup\mathbb{Q}\cup\{i\infty\}$. The rationals $\mathbb{Q}$ have already been identified with the attractive RG fixed points ($\oplus$ in Fig. 1) corresponding to Hall plateaus, while the fixed point at $i\infty$, which is also an attractor in this

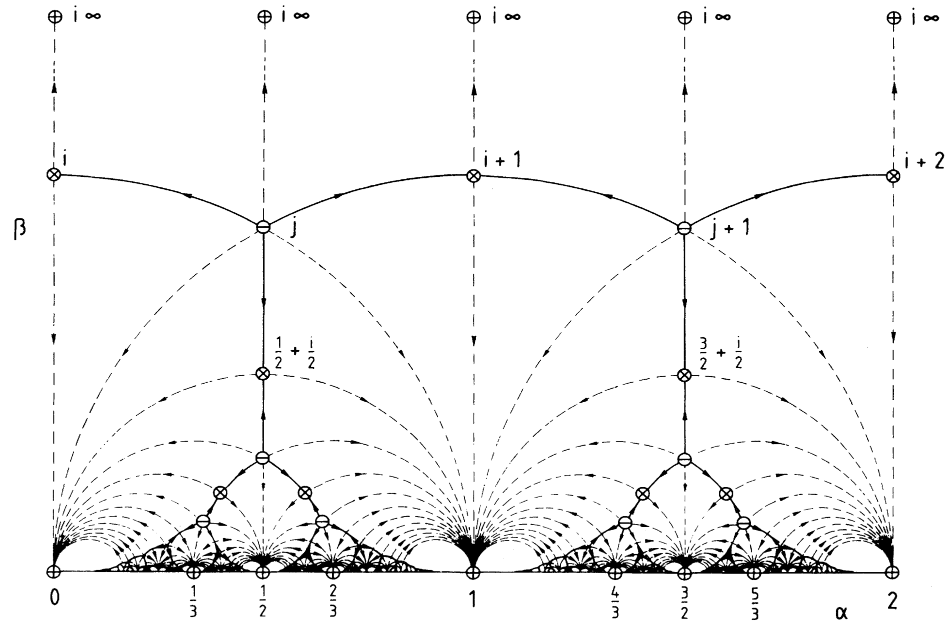

FIG. 1. Modular-invariant phase diagram with simplest assignment of RG flow lines.

case, seems to be some kind of superconductivity fixed point.

There are three distinct scaling diagrams consistent with this fixed-point structure, depending on whether $\gamma(i)$ is a saddle point, a repulsive point, or part of a marginal line of fixed points along the phase boundary. This ambiguity comes about because Kramers-Wannier-type symmetries only pin down the critical *values* of parameters; some additional data are required in order to identify the universality class and thus the scaling exponents encoding the rate at which critical points in parameter space are approached.

If we can, as we shall argue, restrict attention to the class of self-dual [$\Gamma(1)$-invariant] models which interpolate between the coupled Potts models, defined by analytic continuation in p to any positive real p,[13] then the only remaining ambiguity is in the value of p.

Since all fixed points of a given type are mapped into each other under duality, it is sufficient (for $p<2\sqrt{3}$) to consider only the simplest "decoupling fixed point" at $\tau=i$ in order to obtain the scaling exponent of *all* the fixed points, which in the quantum Hall case control the delocalization transitions. Because $Z_p(i)$ in the thermodynamic limit reduces to the critical p-state Potts model, this allows us to read off the values of the critical exponents (ν_α,ν_β) from general results:[14,15]

$$\nu_\alpha=2-2\nu_\beta,\quad \nu_\beta=\frac{2-y}{3(1-y)}\ , \tag{3}$$

where y parametrizes the distance p is away from the marginal (Ising) value $p=2$:

$$p-2=2\cos(\pi y)\ . \tag{4}$$

As shown by Dotsenko and Fateev,[16] this result is but one of many which follow directly from the conformal symmetry that appears at critical points of statistical models. Many of these fall into a sequence of minimal models, labeled by an integer $m=1,2,3,\ldots$, which are completely determined by a finite number of primary fields $\phi_{r,s}$, whose exact scaling dimensions can be determined algebraically from the Kac formula:

$$h_{r,s}=\frac{[rm-s(m+1)]^2-1}{4m(m+1)}\ . \tag{5}$$

Some of the minimal models, with $m=1,2,3,5,\infty$, can be identified with the $p=0,1,2,3,4$ Potts models. The analytical continuation to noninteger values of p should clearly proceed from the more general result (4), which can be derived from a Coulomb-gas formulation. It is not immediately obvious that these conformal constructions can be extended to noninteger values of m, or indeed integer $m<3$, because they contain operators with negative scaling dimension. However, this is not the case for the operator subalgebra containing $\{\phi_{1,n}\}$, which Dotsenko and Fateev conjectured corresponds to the thermal operators; e.g., $\phi_{1,2}$ is known to be the energy operator for both $p=2$ (Ising) and $p=3$ (3-state Potts), because in these cases there is only a finite number of primary fields—the theory is exactly soluble—so that the identification is unambiguous. The complete agreement with previous calculations and conjectures of scaling exponents also for the percolation ($p=1$) and polymer ($p=0$) problem leaves little doubt that this identification is correct.

The anomalous (conformal) dimension of the nth thermal exponent of the mth minimal model is

$$2h_{1,n+1}=\frac{n^2+ny}{2-y}\ , \tag{6}$$

where $y=2/(m+1)$, in agreement with (4). The total scaling dimension is $2-2h_{1,n+1}=\nu_\beta^{-1}(n)$, which for the leading thermal exponent ($n=1$) reduces to $\nu_\beta=2m/3(m-1)$, which is (3) for integer values of p.

Using (3) we can determine the nature of the fixed points for various values of p. The possibilities are evident from Fig. 2 where the critical exponents are plotted as functions of p, and we find

$0 \le p < 2$: The phase and flow diagram is given by Fig. 1.

$p = 2$: The tree in Fig. 1 is marginal (Baxter model).

$2 < p \le 2\sqrt{3}$: The arrows on the tree in Fig. 1 are reversed. (7)

$2\sqrt{3} < p \le 4$: The phases in Fig. 1 are separated by a Coulomb phase.

$p > 4$: No second-order transition.

In short, provided that we are forced into this class of effective-field theories by the microphysics of the quantum Hall system, the critical properties of the model are completely encoded in the value of p, and we are able to determine the scaling exponents of the theory.

It is the purpose of this paper to argue that a realistic treatment of the localization problem in the quantum Hall system leads to a self-dual effective-field theory which can be mapped onto $Z_{p=1}(\tau,\bar{\tau})$. From (3) we see that in this model the delocalization fixed points $\gamma(i)$ are saddle points, as shown in Fig. 1, with critical exponents $(\nu_\alpha,\nu_\beta)=(-\frac{2}{3},\frac{4}{3})$.

III. CONDUCTIVITIES IN CHERN-SIMONS THEORIES

In this section we will review the description of anyonic states in terms of Chern-Simons (CS) theories and discuss the recent work of Kivelson, Lee, and Zhang,[5] who construct an effective-field-theory action describing the propagation of anyonic states and use it to motivate the structure of a global phase diagram for the quantum Hall system. We will demonstrate that what they call "The Law of Corresponding States" corresponds to a subgroup of the full modular group discussed above. Anticipating the results of Secs. V and VI, we will argue that it is the full modular group which is relevant and show how it can emerge from the effective theory.

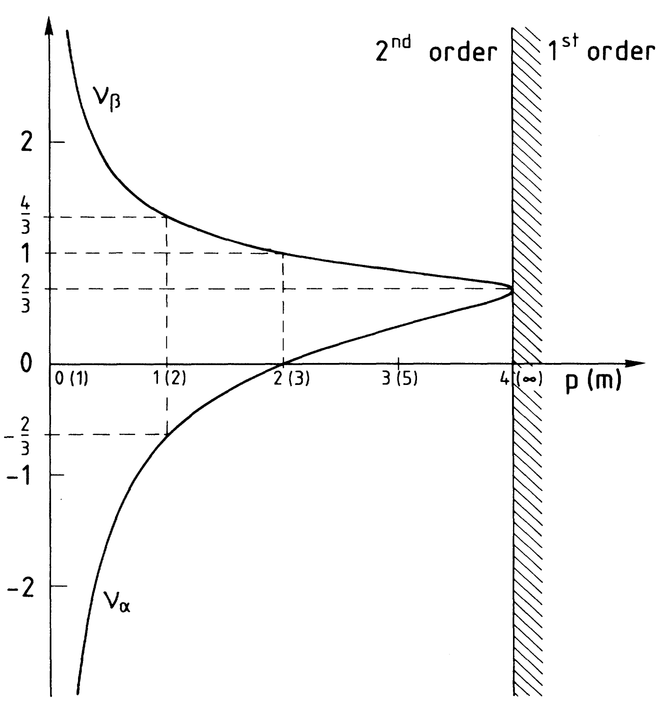

FIG. 2. The critical exponents ν_α and ν_β as functions of p or m.

If we are to describe conductances in terms of the propagation of charge carriers in a random potential, ignoring Coulomb effects, then it must be the nonlocal (anyonic) superpositions of electrons that are the appropriate degrees of freedom. As argued by Laughlin,[17] in this case the effect of the Coulomb forces between electrons is largely taken care of by the construction of the anyonic state, and the choice of the appropriate state needed to cancel the background positive charge distribution. In this case the Green functions of interest describe the propagation of anyons, and the field-theoretic representation of the Green function should involve a path integral over anyonic fields.

The macroscopic properties of these states may be described by an effective Landau-Ginzburg (LG) field theory and we will use this description when determining the conductances. The anyonic state may be bosonized and represented by a complex scalar field ϕ with the Chern-Simons term:

$$L_{\text{any}} = \phi^\dagger i(\partial_0 + ia_0 - ieA_0)\phi + \frac{1}{2M}\phi^\dagger(\partial_i + ia_i - ieA_i)^2\phi + \frac{e^2}{4\theta}\epsilon_{\mu\nu\lambda}a^\mu\partial^\nu a^\lambda \,. \quad (8)$$

This Lagrangian describes anyons obeying θ statistics, i.e., the wave-function changes by a phase θ under the interchange of particles. Under certain plausible assumptions it was shown in Ref. 18 that when θ/π is an odd integer this Lagrangian describes spin-polarized electrons confined to two dimensions in an external transverse magnetic field, neglecting the Coulomb interaction. The Hall conductances may be calculated by applying an external scalar potential A_0 with $\partial_i A_0 = -E_i$. Using the equation of motion this gives[18]

$$\sigma_{ij}E^j = \frac{e^2}{2\theta}\epsilon_{ij}E^j \,. \quad (9)$$

Thus the longitudinal conductance vanishes and the transverse conductance is given by $e^2/2\theta$. The Lagrangian L_{any} therefore describes the behavior to be expected *in the region of the plateaus.*

When θ/π is an odd integer the vacuum state corresponds to the partially filled Laughlin states, and the excitations about these vacua are anyonic with fractional charge and statistics. The presence of such states leads naturally to the generation of the Halperin-Haldane hierarchy of fractional plateaus. Thus, the Chern-Simons theory gives a plausible explanation for the existence of the plateaus in σ_{xy} with odd denominator values.

Recent work[5,19] has attempted to extend this description of the quantum Hall system into the region of transition between levels, by including the effects of fluctuations of the CS fields about their classical values. The result of this analysis is to generate an effective action for the statistical gauge field a_μ and the electromagnetic field A_μ of the form

$$S_{\text{eff}}=\int d^2z\,dt\{-\tfrac{1}{2}\sigma^*_{xy}\epsilon_{\mu\nu\lambda}a_\mu\partial_\nu a_\lambda +\tfrac{1}{2}f_{01}\pi_1 f_{01}+\tfrac{1}{2}f_{12}\pi_2 f_{12} -i\pi_3\epsilon_{\mu\nu\lambda}(\eta A_\mu-a_\mu)f_{\nu\lambda}\}\ , \tag{10}$$

where σ^*_{xy} (called s_{xy} in Ref. 5) is the plateau value of the Hall conductivity in natural units (e^2/h), $f_{\mu\nu}=\partial_\mu(\eta A_\nu-a_\nu)-\partial_\nu(\eta A_\mu-a_\mu)$ and π_1, π_2, and π_3 are space-time functions describing the linear-response properties of the CS bosons.

Since the notation used in Ref. 5 obscures the simple geometrical content of their results, we shall transcribe their formulas into the language introduced above in Sec. II, which is also more appropriate for examining the RG structure of the theory.

In a given *phase,* which is uniquely labeled by a stable fixed point or plateau value $\sigma^*\equiv\sigma_\oplus=\sigma^*_{xy}+i\sigma^*_{xx}=\sigma^*_{xy}=p/q$ of the conductivity, the total conductivity $\sigma=\sigma_{xy}+i\sigma_{xx}$ away from criticality is changed from the fixed-point value σ^* by the excitation of anyonic quasiparticles, which contribute to the transport of charge through the system. This anyonic contribution $\sigma^a=\sigma^a_{xy}+i\sigma^a_{xx}$ to the transport tensor $(\sigma=\sigma^*+\sigma^a)$ can be calculated[5] from the effective action (10) by integrating out the statistical Chern-Simons-gauge field a_μ, and then taking the static limit. In this limit the linear-response functions reduce to the *bosonic* conductivities:

$$\sigma^b_{xx}=\lim_{\omega\to 0}\omega\pi_1(\mathbf{q}=0,\omega),\quad \sigma^b_{xy}=\lim_{\omega\to 0}\pi_3(\mathbf{q}=0,\omega)\ , \tag{11}$$

where ω and $\mathbf{q}$ are the energy and three momentum, respectively. The result of this calculation is therefore parametrized by $\sigma^b=\sigma^b_{xy}+i\sigma^b_{xx}$:

$$\sigma^a=\frac{\theta^2}{\eta^2|\rho|^2}(\sigma^b-\theta|\sigma^b|^2)\ , \tag{12}$$

where the total resistivity $\rho=\rho^*+\rho^a$ is determined by the strikingly simple relation

$$\rho^a=\left[\frac{\theta}{\eta}\right]^2\sigma^b\ . \tag{13}$$

This is equivalent to the results reported in Ref. 5 (up to a sign error in Ref. 5), since the resistivity tensor ρ_{ij} by definition is the inverse of the conductivity tensor σ_{ij} $[\rho\equiv S(\sigma)=-1/\sigma]$, and $\sigma^*=\sigma^*_{xy}=\eta^2/\theta$.[18]

The so-called "Law of Corresponding States" is now easily extracted from the above results. Note first that under the transformation $1/\nu\to 1/\nu+2$, where $\nu=\pi/\theta$ is the filling factor, θ changes by 2π. The statistics of the states described by the CS theory will therefore be unchanged under this transformation, and it is reasonable to expect that the anyonic transport properties, and in particular the bosonic conductivities to which they are related, are unchanged. The propagating anyonic states appear as finite-energy vortex solutions about the vacuum labeled by θ and carry charge $\eta_a=\pm\pi e/\theta$. The connection between the statistics and charge of the anyonic states is given by $\theta_a=\eta_a\Phi/2$, where $\Phi=2\pi/e$ is the flux quantum of the vortex. In short, we see from (13) that ρ^a should not change under this transformation, so that $\rho'-\rho=\rho'^*-\rho^*$. Since $\rho^*=\rho^*_{xy}=-\theta/\eta^2$ we see that $\rho\to\rho+2$ under the transformation $1/\nu\to 1/\nu+2$, provided that σ^b does not change. The rest of "The Law" is obtained in a similar manner: time-reversal symmetry $\sigma(1-\nu)=1-\sigma(\nu)$ follows if $\sigma^b(1-\nu)=J[\sigma^b(\nu)]\equiv-\bar{\sigma}^b(\nu)$ (J is *not* a modular transformation; see below), and since we expect the physics of higher Landau levels to be the same, we should find $\sigma(\nu+1)=\sigma(\nu)+1$, which indeed is true if $\sigma^b(\nu+1)=\sigma^b(\nu)+1$.

The authors of Ref. 5 use this to constrain the global phase diagram. As we demonstrate below these transformations just generate a subgroup of the modular group discussed above in Sec. II. Note that it has been derived making (reasonable) assumptions about the transformations of the bosonic conductivities describing the transport properties of the anyonic states. Ideally these properties should be derived from the properties of the system itself, and we will attempt to do so in the following sections.

Given this appealing physical picture for the origin of the subgroup of the modular group, it is reasonable to ask what would be the physics corresponding to the full modular group. The difference in going to the full modular group is simply that the first transformation should be extended to include the transformation $1/\nu\to 1/\nu+1$. This generates phases with *different* statistics; for example, starting with states obeying fermionic statistics this transformation will give states obeying bosonic statistics. The difference between the phase diagrams shown in Figs. 1 and 3 is just the addition of these phases. In the derivation of the modular transformations in the following sections we will argue that the full phase diagram is relevant and, anticipating this result, we consider its implications in the context of the effective CS theory discussed here.

The important difference in the derivation presented in the following sections is that the parameter for the effective Lagrangian, which becomes the order parameter of the phase transformation when the advanced and retarded sectors are related to each other, is *bilinear* in the CS fields. As a result, changing θ by π in the underlying CS theory leaves the statistics of the "order parameter" invariant. Since the phase structure of the theory is determined by the effective Lagrangian describing this order parameter, the invariance naturally appears in the phase structure of the theory, and therefore gives rise to the full modular group. It is, of course, entirely consistent with (13), given the appropriate invariance of the bosonic conductivities. In particular, the full modular group emerges if we have, in addition to the transforma-

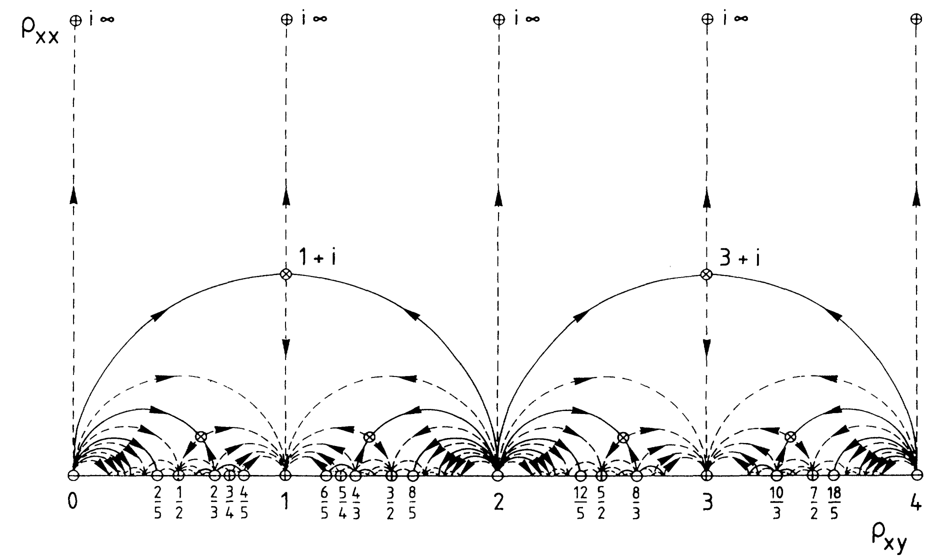

FIG. 3. $\Gamma_W(2)$-invariant phase and flow diagram.

tion properties given above, $\sigma(\nu')=\sigma(\nu)$ for $1/\nu'=1/\nu+1$, which is true if $\sigma^b(\nu')=\sigma^b(\nu)$.

The physical interpretation of these new phases is more problematic. Such phases, which are related to a state with bosonic statistics, may be expected to arise in a system of fermions confined to a plane provided the spin-wave function is antisymmetric, i.e., in states which are not spin polarized. However, this may not be the correct interpretation here. Rather, it is possible that pairing of the fundamental fermion states occurs, giving effective bilinear fields with bosonic properties, and these in turn give rise to the new phases in the effective Lagrangian for the "order parameter." The question whether these phases can be excited depends on the pairing energy and the associated energy of the phase. This is not determined simply from the modular symmetry and requires detailed dynamical information about the anyonic excitation energies which lies beyond the scope of the present work.

The topology, but not the detailed geometry, of the phase diagram (no flow or fixed-point structure was suggested) in Ref. 5 is in fact completely "contained" in the modular-invariant phase and flow diagram proposed in Ref. 1, in the sense that it is determined by one of the simplest subgroups of the modular group $\Gamma(1)$.

The key to identifying this subgroup is to note that whatever else it does, it must map fixed points of a given type into each other. The full modular group maps any rational number into any other, hence it cannot possibly distinguish odd-denominator fractions from even-denominator ones.

If only attractive fixed points (Hall plateaus) $\sigma_\oplus=p/q$ with odd q's are desired, then the subgroup must preserve the parity of q. The group $\Gamma_T(2)$ generated by T and ST^2S, which is obviously contained in $\Gamma(1)$ because any string of these generators is a string of $\Gamma(1)$ generators (S and T), does precisely this: it maps odd (even) denominators to odd (even) denominators, with no restriction on the numerators.

Notice that (the real part of) $\Gamma_T(2)$ is the group implicitly assumed in the so-called "hierarchy generating mechanism,"[17,20,21] which was invented in order to account for the observed fractions (see also Ref. 22). To see this recall that if a ground state with "filling factor" ν appears, then the particle-hole conjugate state with filling factor $1-\nu$, as well as the quasiparticle condensate with filling factor $\nu/(2\nu+1)$, should also be ground states of the quantum Hall system. Since ν is essentially the Hall conductivity on the plateaus, we see that these are fractional linear transformations on σ restricted to act only on the real axis. The first transformation is $TJ=JT$, where $J:\sigma\to-\bar{\sigma}$ is a so-called "outer automorphism" of $\Gamma(1)$. J is in fact the only automorphism of $\Gamma(1)$ not in $\Gamma(1)$, and since J is rather trivial we will continue to suppress it here. The other transformation is the inverse of ST^2S, which together with T generates $\Gamma_T(2)$.

Since the hierarchy generating transformations discussed above played a central role in the work of Ref. 5, it is perhaps not surprising that their diagram appears to have a similar topology to the exact $\Gamma_T(2)$-invariant diagram derived in Ref. 23, but the comparison is not immediate because they chose to work with resistivities rather than conductivities. Since $\rho\equiv\rho_{xy}+i\rho_{xx}=S(\sigma)$, and S is in $\Gamma(1)$, there is no distinction between the phase diagrams for ρ and σ in the fully modular-invariant case. However, since S is *not* in $\Gamma_T(2)$, the phase diagram in resistivity space is in this case different from the phase diagram in conductivity space. The result of S transforming the conductivity phase and flow diagram constructed in Ref. 23 is shown in Fig. 3.

The most economical way of deriving this diagram is to find the group $\Gamma_W(2)$ which acts on the image of σ under S in the same way that S acts on the image of σ under $\Gamma_T(2)$, i.e., to find the S conjugate of $\Gamma_T(2)$, $\Gamma_W(2)=S\Gamma_T(2)S^{-1}$. It is sufficient to conjugate the generators, and using the fundamental identities $S^2=(ST)^3=1$ we find that $\Gamma_W(2)$ is generated by T^2 and $W=TST$. This is sufficient to figure out the fixed-point structure, from which the phase diagram follows. As was the case with $\Gamma_T(2)$,[23] all $\mathbb{E}_3$ fixed points have disappeared, and only some of the $\mathbb{E}_2$ fixed points remain, including the ones at $(2n+1)+i$. Furthermore, $\Gamma_W(2)$ preserves the parity of the *numerator* of the parabolic fixed points, as it should because $S(p/q)=-q/p$ (i.e., an odd-denominator $\sigma_\oplus$ fixed point is an odd-numerator $\rho_\oplus$ fixed point), so that we can consistently choose all odd-numerator fractions to be attractive fixed points and even-numerator fractions to be repulsive fixed points. This completely fixes the phase and flow diagram as shown in Fig. 3.

Figure 3 is to be compared with the resistivity phase diagram presented in Ref. 5. Clearly, the few phases included in Ref. 5 agree with the topology of the exact infinite hierarchy of phases exhibited in Fig. 3. The geometry is however somewhat different. This is presumably due to an arbitrary "normalization" of the phases in Ref. 5. No such freedom is left in our construction—once the group has been fixed the full nonperturbative structure of the phase diagram is rigidly fixed. In particular, the predictions for the location of delocalization fixed points which follow from this are completely falsifiable—they cannot be changed without changing the entire global structure of the diagram, i.e., the discrete symmetry group. Thus these simple symmetry *Ansätze* lead to a surprisingly strong and apparently successful phenomenology of the quantum Hall system. However, current scaling experiments do not appear to be good enough to distinguish between the groups discussed above.

IV. LOCALIZATION OF FERMIONS

In this section we review the field-theoretic approach to localization of ordinary fermions in a way that is appropriate to a discussion of localization in the quantum Hall system.

We are interested in the conductance properties of a two-dimensional system of electrons in a strong transverse magnetic field in the presence of impurities. The effect of the impurities is to provide a reservoir of localized states with energies different from the Landau level, which allow the plateaus to develop. Increasing the magnetic field leads to a reduction in the cyclotron radius causing the Fermi level to fall as localized states drop into the Landau level. While the Fermi level is above the Landau level this change does not alter the transverse conductivity, since the localized states do not conduct, and for the same reason the diagonal (dissipative) conductivity remains zero. As the Fermi energy crosses the Landau level the extended state gives a nonzero diagonal conductivity while the transverse conductivity changes continuously to the value at the plateau corresponding to the next Landau level below the Fermi surface.

In order to describe this system we start with the standard microscopic field-theoretic treatment of electrodynamics in the presence of impurities in two dimensions. We wish to describe the response functions (conductivities) in a field-theoretic treatment of charge transport. In the case of electrons, the diagonal (dissipative) conductivity is determined by the electron Green function computed in a random potential, with some prescribed statistical properties to describe the spreading of the energy level due to impurities in the system. In addition to showing how σ_{xx} emerges as an effective parameter in the macroscopic theory from the underlying microphysics of the semiconductor, the purpose of the following formalism is also to discover how to express the effective action parametrized by σ_{xx} in terms of composite operators Q chosen so that the dissipative conductivity is determined by the propagator of Q. These are given by a "Hubbard-Stratonovich" transformation, which in this case turns out to be a bilinear $Q_{+-}=:\psi_+\psi_-:$ in the advanced and retarded electron fields $\psi_\pm$. Levine, Libby, and Pruisken[24] have given a heuristic interpretation of this in terms of the phase coherence of the charge carrier as it propagates through a noisy environment and strong magnetic field. They argue that localization is related to the destruction of phase coherence between the advanced and retarded propagators, and that the extended states correspond to topologically nontrivial field configurations that restore this phase coherence at the center of the Landau band.

The "advanced" (+) and "retarded" (−) propagators $G_V^\pm=G(z,z';V;E\pm i\eta)$ [from $z=(x,y)$ to $z'=(x',y')$] at fixed energy E and impurity potential V are solutions of the time-independent Schrödinger equation:

$$(\nabla^2+E\pm i\eta-V)G_V^\pm(z,z')=-\delta(z-z') . \tag{14}$$

$G_V^\pm$ describes the propagation of electrons in an impurity potential, *ignoring the interelectron Coulomb interaction.* This should be a reasonable approximation for a *filled* Landau level in the quantum Hall system, since in this case the cyclotron energy is much greater than the Coulomb energy between electrons neutralized by a uniform background electric charge distribution. In the case of partially filled levels, appropriate to the fractional quantum Hall effect, it has been argued that the system may be described by (nonlocal) anyonic states which arise as "the least interacting" collective modes in the electron basis. In this case the conductivities will be determined by the propagation of anyonic states in a random potential, this time ignoring the Coulomb interaction between the *anyonic* charge carriers. In the next section we will discuss the fractional case; here we discuss only integer levels in which case $G_V^\pm$ are propagators of fermionic fields.

The conductivities σ_{xx} and σ_{xy} being physical observables, they are related to the *square* of the Green functions. They are most simply expressed in terms of the "diffusion probability" $\overline{G_V^+G_V^-}\equiv\int\mathcal{D}V\,P[V]G_V^+G_V^-$ (through a sample of unit area) as follows:[25]

$$\sigma_{xx}=\lim_{\eta\to0}\eta^2\int d^2z\,d^2z'(x-x')^2\overline{G_V^+(z,z')G_V^-(z',z)}\ ,$$
$$\sigma_{xy}=i\lim_{\eta\to0}\eta^2\int d^2z\,d^2z'(xy'-x'y)^2\overline{G_V^+(z,z')G_V^-(z',z)}\ . \tag{15}$$

Note that our discussion deals with time-independent solutions so that henceforth only the two-dimensional spatial variables are displayed.

The field-theoretical treatment of these conductivities starts with the functional integral representation of the time-independent (fixed energy) Green functions:

$$G_V^{\pm}(z,z')\equiv\frac{1}{Z_V^{\pm}}\int\mathcal{D}\bar{\psi}_{\pm}\mathcal{D}\psi_{\pm}e^{-S[\bar{\psi}_{\pm},\psi_{\pm};V]}\bar{\psi}_{\pm}(z)\psi_{\pm}(z')\ ,\tag{16}$$

where Z_V is the generating functional and S is the Schrödinger action:

$$S[\bar{\psi}_{\pm},\psi_{\pm};V]=-\tfrac{1}{2}\int d^2z\ \bar{\psi}_{\pm}(\nabla^2+E\pm i\eta-V)\psi_{\pm}\ .\tag{17}$$

The functional average over the impurity potential V involves the distribution $P[V]$, which for convenience we take to be Gaussian,

$$P[V]\equiv\exp\left\{-\frac{1}{2\gamma^2}\int d^2z\ V^2(z)\right\}\ ,\tag{18}$$

but the results should not depend on the precise form of $P[V]$.

Here we have chosen to write the functional integral in terms of fermion fields ψ since, as we will discuss, these are the most appropriate for describing the symmetries associated with the propagation of electrons.

The standard techniques to compute the normalization in (16) involve the replica trick or the supersymmetry trick. In this paper we only use the former, but in view of the difficulties that may be encountered in the analytic continuation of the replica index, it would be worthwhile at some stage to check the results using the supersymmetry technique. Note that not all of our results rely on the validity of the replica trick; to determine the symmetries we do not need the overall normalization of the Green functions.

In the replica trick the normalization factor is eliminated at the price of *replicating* each field ($\psi_{\pm}$) $n_{\pm}$ times, and taking the limits $n_{\pm}\rightarrow 0$ at the end of the calculation. If the original action was Gaussian in V, then so is the replicated one, and we are able to perform the average over impurities. With the replicas labeled by $a=1,2,\ldots,n_{\pm}$, the *normalized* Green functions are given by

$$G_V^{\pm}(z,z')=\lim_{n_{\pm}\rightarrow 0}\int\prod_{a=1}^{n_{\pm}}\mathcal{D}\bar{\psi}^a_{\pm}\mathcal{D}\psi^a_{\pm}e^{-S[\bar{\psi}^a_{\pm},\psi^a_{\pm};V]}\times\bar{\psi}^a_{\pm}(z)\psi^a_{\pm}(z')\ .\tag{19}$$

In the following sums and products over repeated replica indices (early latinos) will usually be suppressed. The diffusion propagator is now determined using (18) and (19):

$$\overline{G_V^{+}(z,z')G_V^{-}(z,z')}=\lim_{n_{\pm}\rightarrow 0}\int\mathcal{D}\bar{\psi}^a_{+}\mathcal{D}\psi^a_{+}\mathcal{D}\bar{\psi}^b_{-}\mathcal{D}\psi^b_{-}e^{-\bar{S}[\bar{\psi},\psi]}\times\bar{\psi}^a_{+}(z)\psi^a_{+}(z')\bar{\psi}^b_{-}(z)\psi^b_{-}(z')\ ,\tag{20}$$

with

$$\bar{S}[\bar{\psi},\psi]=-\tfrac{1}{2}\int d^2z\{\bar{\psi}^a_{+}(\nabla^2+E)\psi^a_{+}+\bar{\psi}^a_{-}(\nabla^2+E)\psi^a_{-}+2i\eta(\bar{\psi}^a_{+}\psi^a_{+}-\bar{\psi}^a_{-}\psi^a_{-})+\frac{\gamma^2}{4}(\bar{\psi}^a_{+}\psi^a_{+}+\bar{\psi}^a_{-}\psi^a_{-})^2\}\ .\tag{21}$$

When $\eta\rightarrow 0$, (20) is invariant under a $U(n_{+}+n_{-})$ symmetry which rotates the full multiplet of fields (ψ^a_{+},ψ^a_{-}). The (small) regulator η explicitly breaks this group to $U(n_{+})\times U(n_{-})$, meaning that (21) is only invariant under independent rotations on the advanced and retarded replicas separately. The importance of these symmetries for the theory of electronic transport was first pointed out by Wegner.[6] In more than two dimensions the quartic term in (21) causes a spontaneous breakdown of the $U(n_{+}+n_{-})$ symmetry to $U(n_{+})\times U(n_{-})$. The Goldstone modes of the associated coset space $U(n_{+}+n_{-})/U(n_{+})\times U(n_{-})$ are the massless states responsible for delocalization and a nonvanishing σ_{xx}. In two dimensions we know from the Coleman-Mermin-Wagner (CMW) theorem that there are no Goldstone phases, so that *generically* there are no delocalized states in low dimensions. In other words, σ_{xx} will typically vanish. However, the effect of the last term of (20) is still to give a mass to the fields which are not in $U(n_{+}+n_{-})/U(n_{+})\times U(n_{-})$. The remaining modes may *occasionally* become massless, thus giving rise to extended states, but in order not to violate the CMW theorem this can happen only at isolated points in parameter space.

In order to quantify this effect it is useful to change variables to the composite fields $:\bar{\psi}^a_{\pm}\psi^b_{\pm}:$. This is done by introducing a bosonic Hermitian matrix of Lagrange multiplier fields $Q^{ab}_{\pm\pm}$ and multiplying (20) by the Gaussian factor

$$\int\mathcal{D}Q\ \exp\tfrac{1}{2}\{Q^{ab}\cdot Q^{ab}-\gamma\bar{\psi}^a\cdot Q^{ab}\cdot\psi^b\}\ ,\tag{22}$$

where the dot denotes matrix multiplication in the two-dimensional "Wegner space" labeled by the advanced (+) and retarded (−) index. The advantages of changing variables to the Q's is that the Green function of Q determines the diffusion propagator, and hence the conductances (15). Thus, if in the long distance limit there are massless states in the field theory describing the Q's, there will be extended states in the system of interest corresponding to a nonzero value for the linear conductance. The Green functions of the Q's may be found from the expression for the partition function given entirely in terms of the Q's, which can be obtained by integrating over the fermion fields $\psi_{\pm}$ in (20) extended by the factor (22), giving[25]

$$Z[J]=\lim_{n_{\pm}\rightarrow 0}\int\mathcal{D}Q\ \exp\tfrac{1}{2}\left\{\mathrm{Tr}\ln A(Q)+\int d^2z(Q^{ab}\cdot Q^{ab}-J^{ab}\cdot Q^{ab})\right\}.\tag{23}$$

Here the J's are sources for the Q fields, and the matrix

$$A=\begin{bmatrix}\nabla^2+E+i\eta+\gamma Q_{++} & -\gamma Q_{+-}\\ -\gamma Q_{-+} & \nabla^2+E-i\eta+\gamma Q_{--}\end{bmatrix}\tag{24}$$

determines the effective action in the Q basis.

The Q Lagrangian contains terms linear in Q and these induce a vacuum expectation value (vev) for Q, which can be determined from the condition

$$\frac{\delta}{\delta Q}\left\{\mathrm{Tr}\ln A(Q)+\int d^2zQ^{ab}\cdot Q^{ab}\right\}\Bigg|_{Q=\langle Q\rangle}=0\ , \quad (25)$$

giving

$$\langle Q_{+-}\rangle=\langle Q_{-+}\rangle=0\ ,$$
$$\langle Q_{\pm\pm}\rangle=-\frac{\gamma}{2}(\nabla^2+E\pm i\eta+\gamma\langle Q_{\pm\pm}\rangle)^{-1}_{zz}=\frac{\gamma}{2}G_0^{\pm}(z,z)\ . \quad (26)$$

Next we expand (23) in fluctuations $\hat{Q}\equiv Q-\langle Q\rangle$ about this minimum. This gives the quadratic terms in $\hat{Q}_{+-}$:

$$\tfrac{1}{4}\int d^2z\, d^2z'\{\hat{Q}_{+-}(z)C_{+-}(z,z')\hat{Q}_{-+}(z') +\hat{Q}_{-+}(z)C_{+-}(z,z')\hat{Q}_{+-}(z')\}\ , \quad (27)$$

where

$$C_{+-}(z,z')=2\delta(z-z')-\gamma^2G_0^{+}(z,z')G_0^{-}(z',z) \quad (28)$$

and

$$G_0^{\pm}(z,z')=-(\nabla^2+E\pm i\eta+\gamma\langle Q_{\pm\pm}\rangle)^{-1}_{xx'} \quad (29)$$

are the advanced and retarded Green functions at tree level. The density of states at tree level is given by [see (26)]

$$2\pi i\overline{\rho_0(E)}=\lim_{\eta\to 0}[\overline{G_0^{+}(0,0)}-\overline{G_0^{-}(0,0)}] =\frac{2}{\gamma}[\langle Q_{++}\rangle-\langle Q_{--}\rangle]\ . \quad (30)$$

After shifting to the fields $\hat{Q}$, those fields not in the coset space $U(n_++n_-)/U(n_+)\times U(n_-)$ acquire a mass. In order to project onto the remaining massless fields it is convenient to change variables by writing $Q(z)=T^{-1}(z)P(z)T(z)$, where $T(z)\in U(n_++n_-)$ and $P(z)$ is block diagonal in the advanced and retarded subspaces. The advantage of this is that the fluctuations of the P fields about their vev's are massive and hence do not contribute to the critical behavior. They cannot, therefore, appear in the effective action governing the properties of the system at large length scales.

The fields in the coset space $U(n_++n_-)/U(n_+)\times U(n_-)$ all belong to the T fields, which can be written as

$$T(z)=\exp\{i\phi_a(z)J^a\}\ , \quad (31)$$

where J^a are generators of $U(n_++n_-)$ not in $U(n_+)\times U(n_-)$ and not commuting with $\langle Q\rangle$. From (30) we see that

$$\hat{Q}=\frac{\pi\gamma i}{2}\overline{\rho(E)}(T^{-1}\tau_{3r}T-\tau_{3r})\ , \quad (32)$$

where τ_{3r} is the diagonal generator of the Wegner symmetry [τ_3 in the unreplicated case of $SU(2)$], which is unbroken by $\langle Q\rangle$. To make contact with Pruisken's analysis[8] it is convenient to define a new matrix field

$$\tilde{Q}(z)=T^{-1}(z)\tau_{3r}T(z)\ , \quad (33)$$

which contains only the light coset fields $T(z)$.

In terms of the $\tilde{Q}$ fields defined in (33) the leading part of the effective action (i.e., second order in derivatives) is

$$L_\sigma=-\frac{1}{8}\sigma^0_{xx}\mathrm{Tr}\partial_\mu\tilde{Q}\partial^\mu\tilde{Q}+\frac{i}{8}\sigma^0_{xy}\mathrm{Tr}\epsilon_{\mu\nu}\tilde{Q}\partial^\mu\tilde{Q}\partial^\nu\tilde{Q}\ . \quad (34)$$

The first term just follows from (27), while the second term is the one identified by Pruisken and coworkers[8,24,26] as necessary to give a nonvanishing σ_{xy}. The form of (34) follows from (15), with the superscript "0" denoting that the vev's have been determined using the zeroth-order theory.

The second term in (34) is a nontrivial topological invariant because it is the Jacobian of the mapping $\tilde{Q}$ from the (compactified) worldsheet S^2 to the coset space $U(n_++n_-)/U(n_+)\times U(n_-)$, whose second homotopy group is $\mathbb{Z}$. This implies that the coefficient σ^0_{xy} must be periodic with period 2π, since otherwise the partition function would not be single valued. The implications of (34) for the transition between levels has been discussed by Affleck.[27] He argues that it governs transitions between integer levels with critical points at $\sigma_{xy}=m+\frac{1}{2}$, $m\in\mathbb{Z}$. The critical theory is identified with the $n\to\infty$ limit of the $SU(n)\times SU(n)$ Wess-Zumino-Witten model, leading to a scaling exponent of $\frac{1}{2}$, which, however, is in disagreement with the recently measured values. As we will argue in the following sections, the inclusion of a more general form for the impurity potential changes this expectation.

V. LOCALIZATION OF ANYONS

We consider now how the discussion of Sec. IV changes when discussing the fractional levels. The obvious difference in describing the fractional levels is that they correspond to partially filled Landau levels, and as discussed in Sec. III, these may be described in a field-theoretic sense by the Chern-Simons theory $L_{\rm any}$ given by (8). Although the form of $L_{\rm any}$ is adequate for describing the transport properties of anyonic states for vanishing σ_{xx}, it is necessary to include the effects of impurity scattering if we are to understand transitions between levels. Thus we modify (8) by adding a LG potential $W(\phi,V)$ with form yet to be determined, but depending on both the LG field ϕ and the impurity potential V.

As we discussed in the preceding section, the analysis of the transport properties of the system proceeds via the identification of the relevant degrees of freedom which are massless at the critical points. In this the Wegner symmetry played a crucial role and so we first discuss the symmetries of (8), allowing for a replica index a and an advanced or retarded index ($\pm$) on the LG field, i.e., $\phi=\phi^a_\pm$. The action needed to determine the diffusion propagator is now [cf. (8)]

$$L^W_{\rm any}=\phi^a_+(\nabla^2-i\partial_0)\phi^a_++\bar{\phi}^a_-(\nabla^2-i\partial_0)\phi^a_- +2i\eta(\bar{\phi}^a_+\phi^a_+-\bar{\phi}^a_-\phi^a_-) +\frac{e^2}{4\theta}\epsilon_{\mu\nu\lambda}(a^{\mu a}_+\partial^\nu a^{\lambda a}_++a^{\mu a}_-\partial^\nu a^{\lambda a}_-) +W(\phi^a_+,V)+W(\phi^a_-,V)\ , \quad (35)$$

where $\partial_0=\partial/\partial_t$, and t is a "time" variable introduced in order to cast the CS Lagrangian into a local form.

It is clear from this equation that the Wegner $U(n_++n_-)$ symmetry does *not* act linearly on the fields $\phi^a_\pm$, because of the auxiliary fields $a^{\mu a}_\pm$. Indeed, once these fields are integrated out to give a (nonlocal) Lagrangian[4] expressed entirely in terms of the LG fields $\phi^a_\pm$, one finds that it contains quartic terms in $\phi^a_\pm$ which are not invariant under the $U(n_++n_-)$ symmetry acting linearly on the $\phi^a_\pm$ basis. To exhibit the linear $U(n_++n_-)$ symmetry it is necessary to reformulate (35) directly in terms of the anyonic fields $\phi^a_{\rm any}$ satisfying "anyonic boundary conditions." For the general case of anyons in the absence of the potential W the equivalent formulation has been given by Semenoff.[28] Of course, the simplest example of this is the case $\theta=\pi$ corresponding to fermionic fields. In this case the action written in terms of the fermionic fields $\psi^a_\pm$ is just the one in (21) which manifests the $U(n_++n_-)$ symmetry acting linearly on the $\psi^a_\pm$ basis. For the general anyonic case the anyonic fields $\phi^a_{\pm{\rm any}}$ are nonlinearly related to the LG fields $\phi^a_\pm$ appearing in (35). Since the Wegner symmetry acts linearly on the $\psi^a_\pm$ it is clear that it acts nonlinearly on the LG fields.

So far we have not discussed the description of impurities which follows from the form of W. In the case of fermions the analysis leading to (34) relied on the assumption that $W(\psi^a_+,V)+W(\psi^a_-,V)$ was $U(n_++n_-)$ invariant. Indeed the choice of $W[\psi^a_\pm(z),V]=V\bar\psi^a_\pm(z)\psi^a_\pm(z)$ clearly leads to such a symmetric form. This choice corresponds to the (reasonable) assumption that electrons scatter in a pointlike manner from the impurities. However, the assumption seems less viable for extended anyonic states of fractional spin, since they describe nonlocal superpositions of electrons. Only if the impurities scatter via long-range potentials with range larger than the size of the anyons is it a reasonable approximation to take the equivalent pointlike coupling $W[\phi^a_{\pm{\rm any}}(z),V]=V\bar\phi^a_{\pm{\rm any}}(z)\phi^a_{\pm{\rm any}}(z)$ of the anyon to the impurity potential V. In the experimental configurations it *is* often the case that the impurities are long range, so we will continue to take such a $U(n_++n_-)$-invariant form for the coupling as the *dominant* effect of noise, but we will allow for the presence of additional terms to take account of the possibility of noncoherent scattering of the anyonic states from the impurities. With the assumption that the $U(n_++n_-)$-invariant noise term is dominant, the identification of the light degrees of freedom appropriate near criticality is the same as discussed in the fermionic case, namely the coset fields in $U(n_++n_-)/U(n_+)\times U(n_-)$. We first construct the effective Lagrangian in this approximation before considering the important effects of the symmetry-breaking terms.

Following from the identification of symmetries the analysis proceeds in an entirely equivalent way to that presented in Sec. IV. The first step is to change variables to the composite fields $Q^{ab}_{\pm\pm}\equiv:\bar\phi^a_{\pm{\rm any}}\phi^b_{\pm{\rm any}}:$. The effect of the $U(n_++n_-)$-invariant component of $W(\phi^a_{+{\rm any}},V)+W(\phi^a_{-{\rm any}}V)$ is to generate a vev for $Q^{ab}_{\pm\pm}$, and only the fields in the coset space $U(n_++n_-)/U(n_+)\times U(n_-)$ remain massless at tree level. To lowest (second) order in derivatives, the effective action governing the Q fields is again given by (34). However, as we have just discussed, we expect that at some level terms noninvariant under $U(n_++n_-)$ will appear in W. These we will allow for by including the most general perturbation of L_σ consistent with the unbroken symmetries of the problem.

As we discussed in Sec. IV, the massless fluctuations are contained in $\tilde Q=T^{-1}\tau_3T$. Near the critical point only the most relevant operators are significant and these correspond to keeping the leading terms (linear in ϕ) in the expansion of (31). From (32) we see that $\tilde Q$ may, up to an additive constant, be identified with the fields $Q_{+-}\gamma/\pi i\overline{\rho(E)}$. Hence, near a critical point we may replace (34) by

$$L_{\rm kin}+L_{\rm top}\equiv-\sigma^0_{xx}{\rm Tr}\partial_\mu Q_{+-}\partial^\mu Q_{+-}+i\sigma^0_{xy}{\rm Tr}\epsilon_{\mu\nu}\partial^\mu Q_{+-}\partial^\nu Q_{+-}\ . \qquad (36)$$

In writing (36) we have absorbed a constant in the definition of the normalization of the field Q_{+-}. The relative normalization of the terms proportional to σ_{xx} and σ_{xy} is unaffected by this choice. The absolute normalization of the Q fields will be determined by the condition that the correct value for σ_{xy} is obtained on the plateaus from the underlying CS theory.

The result (36) is of the same form as (34) which described the fermionic case. As such it presents a major problem in understanding the fractional effect, for, following Affleck's analysis,[27] we would expect critical points at half integer values only. The reason the effective Lagrangians are the same follows from the fact the symmetries of the initial Lagrangians are the same, leading to the identification of the same massless modes. However, as we emphasized above, the anyonic case may be expected to differ from the fermionic case because the Wegner symmetry is likely to be explicitly broken by noise terms.

The effect of the explicit $U(n_++n_-)$ breaking terms may now be included as the most general perturbation $L_{\rm int}$ of (36):

$$L_{\rm int}=m^2|Q^\alpha_{+-}|^2+m'^2(Q^\alpha_{+-})^2+\lambda|Q^\alpha_{+-}|^4+\lambda'|Q^\alpha_{+-}|^2|Q^\beta_{+-}|^2+\eta(Q^\alpha_{+-})^4+\eta'(Q^\alpha_{+-})^2(Q^\beta_{+-})^2+{\rm h.c.}+\cdots\ , \qquad (37)$$

where $\alpha,\beta=1,2,\ldots,n_+n_0$ are (composite) replica indices, and the dots denote higher dimension terms which are irrelevant operators in the scaling limit. Since $m^2,m'^2\to0$ at criticality, the terms determining the universality class of this action are *quartic* in the massless fields. In short, the effect on the LG action in the Q basis of including anyonic excitation is to allow a more general LG potential which *may* lead to a different universality class from that expected for fermions.

The effective Lagrangian $L_{\rm eff}\equiv L_{\rm kin}+L_{\rm top}+L_{\rm int}$, given by (36) and (37), has been derived largely from symmetry arguments, and applies when there is an approximate $U(n_++n_-)$ symmetry broken to $U(n_+)\times U(n_-)$ by the Q_{++} and Q_{--} vev's induced by the $U(n_++n_-)$-

invariant noise term. Note that the approximation of keeping only the Q_{+-} fields in the effective Lagrangian appropriate for discussing the physics at critical points is reasonable only if the noise terms in (37), which explicitly break the $U(n_+ + n_-)$ symmetry, are small relative to the $U(n_+ + n_-)$-invariant noise terms.

We must now determine the properties of the LG Lagrangian L^W_{any}, whose low-energy degrees of freedom have been encoded in L_{eff}. Even though the terms of L_{int} are initially small, as we approach the critical point they play an important role in determining the universality class to which the effective Lagrangian belongs. In this we are guided by the symmetries of the effective Lagrangian.

We consider first the case in which we do not take the replica limit. As mentioned above the replica limit is not necessary if we merely wish to discuss the symmetries of the effective Lagrangian. In this case the symmetry of L_{eff} is just a $\mathbb{Z}_4$ factor of the third $[U(1)]$ component of the $U(2)$ Wegner symmetry acting on the (unreplicated) advanced and retarded subspaces. This is left unbroken by (37) simply because the most relevant operator at criticality is quartic in the Q fields. Because of this term the LG theory is in the same universality class as two coupled Ising ($p=2$ Potts) models with a permutation symmetry between the two factors. This is just one of the models discussed in Sec. III, and it has the self-dual symmetry leading to the SL(2,$\mathbb{Z}$) discrete symmetry acting on the conductivities which was used in Ref. 1 to relate the integer and fractional effects.

We conclude that, due to the additional explicit symmetry-breaking terms of (37), at criticality the effective theory L_{eff} will be driven in the long-distance limit to one of the coupled spin models L_p, and *not* to the σ-model L_σ in (34). Unlike the latter model, L_{eff} does allow for the fractional levels associated with the CS theory of (8).

We are thus able to make a connection between two types of effective "Landau-Ginzburg" descriptions of the quantum Hall system. One of these is a "mesoscopic" LG description of the plateaus in terms of anyonic states and the associated CS form of the effective Lagrangian. Here we have related the static properties of such a Chern-Simons theory to a "macroscopic" two-dimensional effective "σ-model with potential" type of LG field theory, which is directly parametrized in terms of the conductances and which can be interpreted as being in the universality class of a self-dual spin model. The discrete symmetry of this spin model allows us to determine the phase structure and positions of (some or all) the RG fixed points. This was discussed in Sec. II and in more detail in Ref. 1, where it was shown that these aspects of the model are in good agreement with experiment.

The macroscopic theory derived here complements and extends the analysis of Ref. 5, for we have seen that the RG flow drives the theory at macroscopic scales to a theory in which the transport properties of the anyonic states are determined and the symmetries relating levels are manifest. As we discussed in Sec. III, these symmetries may be understood as a combination of time-reversal or particle-hole transformations, Landau-level translations, and transformations which do not change the statistics of the anyonic state which is playing the role of order parameter. Since the order parameter is bilinear in the original anyonic field, the latter symmetry is larger than would have been encountered in the original Chern-Simons theory and leads to the appearance of new even-denominator phases. The phase structure is determined entirely from the modular symmetry, but symmetry arguments alone do not determine the energetics of the various phases. If the new "bosonic hierarchy" of phases lies higher in energy, e.g., due to a large pairing energy, then they will not be populated and only the odd-denominator phases will appear. However, their existence still plays an important role in determining the position of the fixed points and hence the behavior of the system in the odd phases near these fixed points.[1]

Notice also that not every energetically allowed phase will appear in every experiment. Which plateau is actually observed is an initial value problem: different phases can be "dialed" by setting external "control" parameters. The magnetic field (B) is just one of these. The degree of disorder is another, which is related, by standard semiclassical scattering/diffusion theory, to the value of the dissipative conductivity σ_{xx}. For a given value of disorder (i.e., a given sample), not any value of σ_{xx} can be obtained ("dialed") by varying B, say. Fixing all external parameters gives an "initial value" to the RG flow, which is then generated by changing *only* the scale parameter. Starting with a dirty sample we only have large initial values available for the flow, i.e., we are always in an integer phase and only such plateaus ($\oplus$=integer) are observed. With a cleaner sample we can get down into a fractional phase, and we will see also some fractional plateaus ($\oplus$=fraction) when B is varied. The cleaner the sample the more phases (plateaus) can be accessed, but each plateau becomes narrower. Eventually they disappear completely, as they should because no impurities means no localized states means no plateaus, and the classical result is recovered.

We are also able to use the macroscopic LG description to investigate the critical behavior corresponding to transitions between plateaus and to determine the critical indices. In order to do this it is necessary to determine which is the self-dual theory that is relevant in this case, and to do this we must take the replica limit, which we now discuss.

VI. REPLICA LIMIT, CRITICAL EXPONENTS, AND EXPERIMENTS

As we discussed in Sec. II, if the delocalization exponent is greater than 1, the crossover exponent at the "decoupling" fixed point at $\sigma_\otimes = i$ is negative, and the RG flow diagram is the one shown in Fig. 1. For this particular fixed point the analysis simplifies considerably, because L_{top} is absent from L_{eff}, or equivalently, $\beta=\sigma_{xy}$ is absent from L_p.

We must take the replica limit $n_\pm \to 0$ to find the theory appropriate for determining the scaling behavior of the conductances. In order to determine the universality class of the LG theory L_{eff} we rely on identifying the symmetry. Adding the replica index the symmetry of $L_{\rm kin}$ is $U(n_+ + n_-)$. This is respected by the first term of $L_{\rm int}$ but broken to $\mathbb{Z}_{(n_+ n_-)!} \times \mathbb{Z}_2^{1+n_+ n_-}$ by the additional terms, where the first factor is the permutation symmetry acting on the replicas Q^α_{+-}. The second factor consists of a $\mathbb{Z}_2$ permutation between the real and imaginary components of Q_{+-}, and a $\mathbb{Z}_2^{n_+ n_-}$-invariance under sign changes of Q^α_{+-}. The potential corresponding to (37) has (for a suitable range of the parameters) $p \equiv 2^{n_+ n_-}(n_+ n_-)!$ minima for the real components of Q^α_{+-} and similarly for the imaginary components.

Since the field variables will only take values at these minima in the long-distance limit, it is reasonable to conjecture that the model lies in the universality class of the p-state Potts model. While we are not able at present to justify this conjecture directly, we will show that it does offer an explanation of the measured value of the delocalization exponent. The first point to note is that this theory has a well-defined replica limit, $p \to 1$ when $n_\pm \to 0$, which corresponds to the "classical" or "geometrical" percolation theory with critical exponent $\nu_\beta = \frac{4}{3}$. It is encouraging that this analysis has identified the universality class to which percolation belongs as the one relevant for the quantum Hall system, since this conclusion can also be reached from the semiclassical "percolation picture"[29] in which the extended state corresponds to delocalization of previously localized states as the Fermi surface crosses the percolation threshold where the localized trajectories coalesce. Although $\frac{4}{3}$ is not the measured value of the delocalization exponent ν_β, it has been argued[30] that "quantum percolation" will dominate away from criticality, leading to a critical exponent of $\frac{7}{3}$ which is in good agreement with the experimental results quoted below.

The obvious question is whether such an explanation applies also to the field-theoretic analysis presented here. The idea behind the quantum-percolation calculation is that, away from the percolation threshold, propagation of charge proceeds via percolation through the domains where the localized trajectories coalesce, whose characteristic size is given by the percolation length, together with tunneling between the different domains. Close to a critical magnetic-field strength $B^* = B_\otimes$ the percolation length scales as $(B - B^*)^\nu$, while the tunneling probability scales as $(B - B^*)$, so that the two together scale as $(B - B^*)^{\nu+1}$. Note that the tunneling amplitude is proportional to $\exp\{-\int_{-a}^{a} dr \sqrt{V(r) - E}\}$, where $2a$ is the distance between domains.

To see whether tunneling processes are included in the LG theory discussed in the preceding section, we consider the approximations used in deriving the effective Lagrangian $L_{\rm eff}$. In the derivation of (36) and (37) the contributions from the fields Q_{++} and Q_{--} were dropped, for $L_{\rm eff}$ was constructed under the assumption that their masses, proportional to γ^2, are large. The propagation of these fields leads to correlation functions proportional to $\exp(-m|z - z'|)$, where the mass m of the field is of order γ. Since γ^2 is the width of the Gaussian determining the range of the potential we see that m is of order $\sqrt{V}$, which is just the order of the terms determining the tunneling amplitude. Thus the approximation used in the derivation of the effective Lagrangian amounts to ignoring tunneling effects. This is consistent with the derivation of the percolation exponent just discussed. At criticality this is the correct theory to describe the conductances, and in particular the value of σ_{xx}. However, away from criticality the effects of tunneling must be added, changing the scaling exponent from $\frac{4}{3}$ to $\frac{7}{3}$, as discussed above.

This result can now be compared with experiments. Until very recently the only exponent that could be measured was the temperature exponent κ, which describes how fast the peaks in σ_{xx} (or, equivalently, the slope at the inflection point between plateaus in σ_{xy}) change as the temperature vanishes. In these experiments[31] it was found that κ takes on a universal value $\kappa = 0.42 \pm 0.04$, irrespective of which transition was considered. This exponent is related to the delocalization exponent ν_β through a standard scaling argument[32] that gives $\nu_\beta = p/2\kappa$, where p is the temperature exponent of the inelastic-scattering rate. The value of the exponent p was not known in strong magnetic fields, but assuming that the value $p \simeq 1.1$ for metals in zero field is still valid in the Hall experiment, it was suggested that $\nu_\beta \simeq 1.3 \simeq \frac{4}{3}$, in apparent agreement with the *classical* percolation exponent. More recent experiments,[33] however, have found that the value of κ may vary by up to a factor of 2 between some semiconductor heterostructures, thus effectively killing off the universality hypothesis for κ.

Fortunately, in a remarkable experiment[34] all three exponents appear to have been obtained *independently* by studying a sequence of different Hall bars, which do not differ in shape or composition, but only by an overall scaling factor. The basic idea is that the temperature and sample widths are so small that the inelastic-scattering length is frozen out of the problem (since it exceeds the sample size), so that *the sample width can be interpreted as the RG scale of the problem*. They find that while both κ and p are nonuniversal, the delocalization exponent takes the universal value $\nu_\beta = 2.3 \pm 0.1$, in agreement with the result of numerical simulations[35] for the lowest Landau level, $\nu_\beta = 2.34 \pm 0.04$. Both results are clearly consistent with the (quantum percolation) exponent $\nu_\beta = \frac{7}{3}$ obtained above from the self-dual effective-field theory of anyon localization.

Finally we remark that the authors of Ref. 34 also found that the value of $\sigma_{xx}^{\rm max}$ appears to depend on the random potential present in the sample, apparently in contradiction with the alleged "superuniversality" of the critical behavior. However, this may be just a reflection of the variety of critical points, related by the modular symmetry, which are predicted in the system. Indeed, in Ref. 36 we showed that the observed values of $\sigma_{xx}^{\rm max}$ are consistent with the predictions for the positions of the critical points which follow from the modular symmetry. While this identification is still tentative and requires further experimental investigation, the agreement lends further support to the suggestion that the quantum Hall system displays the full modular symmetry discussed above.

VII. SUMMARY AND CONCLUSIONS

In summary, we have constructed an effective-field theory for the quantum Hall system which embodies both conventional localization and the possibility of fractional statistics. The latter involves a Chern-Simons interaction, while the former is a generalization of localization theory which still exhibits scaling. Because our treatment of noise is somewhat different, and we claim more realistic, than the standard one, we do not recover Wegner's or Pruisken's σ-models, but rather an effective theory which is invariant under "complexified" duality transformations of the conductivities which appear as effective parameters in the model. Without resorting to an explicit computation of instanton-driven RG flows, but instead exploiting the symmetries of the parameter space as well as the conformal symmetry which appears at RG fixed points, we are able to extract a precise prediction for the whole scaling diagram, which contains both fractional and integer phases. Other predictions are the exact location of all fixed points, and "superuniversality" of the scaling exponents. A plausible identification of the theory in the replica limit leads to the value $\frac{7}{3}$ for the critical exponent of the delocalization transition between plateaus in the Hall conductivity. The agreement with available scaling experiments is excellent.

ACKNOWLEDGMENTS

C.A.L. wishes to thank the Norwegian Research Council for Science and Humanities (NAVF) and the British Science and Engineering Research Council (SERC) for financial support. G.G.R. wishes to thank SERC for support. We are grateful to I. Aitchison, J. Chalker, T. Hollowood, R. Nicholas, J. Singleton, R. Stinchcombe, and J. Zuk for useful discussions.

*Present address: Nordita, Blegdamsvej 17, DK-2100 Copenhagen Ø, Denmark.

[1]C. A. Lütken and G. G. Ross, Phys. Rev. B **45**, 11 837 (1992).

[2]J. M. Leinaas and J. Myrheim, Il Nuovo Cimento **37**, 1 (1977).

[3]F. Wilczek, Phys. Rev. Lett. **59**, 957 (1982).

[4]D. P. Arovas, J. R. Schrieffer, and F. Wilczek, Phys. Rev. Lett. **53**, 722 (1984); D. P. Arovas, J. R. Schrieffer, F. Wilczek, and A. Zee, Phys. Nucl. Phys. B **251**, 117 (1985).

[5]S. Kivelson, D.-H. Lee, and S.-C. Zhang, Phys. Rev. B **46**, 2223 (1992).

[6]F. Wegner, Z. Phys. B **35**, 207 (1979); **36**, 209 (1979); **38**, 113 (1980); **51**, 279 (1983).

[7]V. T. Dolgopolov, A. A. Shashkin, N. B. Zhitenev, S. I. Dorozhkin, and K. von Klitzing, Phys. Rev. B **46**, 12 560 (1992).

[8]A. M. M. Pruisken, in *The Quantum Hall Effect,* edited by R. E. Prange and S. M. Girwin, Graduate Texts in Contemporary Physics, 2nd ed. (Springer-Verlag, Berlin, 1990), and references therein.

[9]J. L. Cardy and E. Rabinovici, Nucl. Phys. B **205**, 1 (1982).

[10]J. L. Cardy, Nucl. Phys. B **205**, 17 (1982).

[11]H. A. Kramers and G. H. Wannier, Phys. Rev. **60**, 252 (1941).

[12]R. J. Baxter, Phys. Rev. Lett. **26**, 832 (1971).

[13]C. M. Fortuin and P. W. Kasteleyn, Physica **57**, 536 (1972).

[14]M. P. M. den Nijs, J. Phys. A **12**, 1857 (1979).

[15]J. L. Black and V. J. Emery, Phys. Rev. B **23**, 429 (1981).

[16]V. S. Dotsenko and V. A. Fateev, Nucl. Phys. B **240**, 312 (1984).

[17]R. Laughlin, Phys. Rev. Lett. **50**, 1395 (1983).

[18]S. C. Zhang, T. H. Hansson, and S. Kivelson, Phys. Rev. Lett. **62**, 82 (1989).

[19]D.-H. Lee and S.-C. Zhang, Phys. Rev. Lett. **66**, 1220 (1991).

[20]F. D. M. Haldane, Phys. Rev. Lett. **51**, 605 (1983).

[21]B. Halperin, Phys. Rev. Lett. **52**, 1583 (1984).

[22]A. H. MacDonald, G. C. Aers, and N. W. C. Dharmawardana, Phys. Rev. B **31**, 5529 (1985).

[23]C. A. Lütken (unpublished).

[24]H. Levine, S. Libby, and A. M. M. Pruisken, Phys. Rev. Lett. **51**, 1915 (1983); Nucl. Phys. B **240**, 30 (1984); **240**, 49 (1984); **240**, 71 (1984); A. M. M. Pruisken, *ibid.* **235**, 277 (1984).

[25]A. J. McKane and M. Stone, Ann. Phys. **131**, 36 (1981).

[26]D. E. Khmel'nitzkii, Pis'ma Zh. Eksp. Teor. Fiz. **38**, 454 (1983) [JETP Lett. **38**, 552 (1983)]; H. Levine and S. B. Libby, Phys. Lett. B **150**, 182 (1985); A. M. M. Pruisken, Phys. Rev. B **32**, 2636 (1985).

[27]I. Affleck, Nucl. Phys. B **265**, 409 (1986), and references therein.

[28]G. W. Semenoff, Phys. Rev. Lett. **61**, 517 (1988).

[29]S. A. Trugman, Phys. Rev. B **27**, 7539 (1983).

[30]G. V. Mil'nikov and I. M. Sokolov, Pis'ma Zh. Eksp. Teor. Fiz. **48**, 494 (1988) [JETP Lett. **48**, 536 (1988)].

[31]H. P. Wei, D. C. Tsui, M. A. Paalanen, and A. M. M. Pruisken, Phys. Rev. Lett. **61**, 1294 (1988); L. Engel, H. P. Wei, D. C. Tsui, and M. Shayegan, in *Proceedings of the Eighth International Conference on Electronic Properties of Two-Dimensional Systems, Grenoble, France, 1989* [Surf. Sci. **229**, 13 (1990)].

[32]D. J. Thouless, Phys. Rev. Lett. **39**, 1167 (1977).

[33]S. Koch, R. J. Haug, K. v. Klitzing, and K. Ploog, Phys. Rev. B **43**, 6828 (1991).

[34]S. Koch, R. J. Haug, K. v. Klitzing, and K. Ploog, Phys. Rev. Lett. **67**, 883 (1991).

[35]J. T. Chalker and P. D. Coddington, J. Phys. C **21**, 2665 (1988); B. Huckestein and B. Kramer, Phys. Rev. Lett. **64**, 1437 (1990).

[36]C. A. Lütken and G. G. Ross (unpublished).