(Mis-)handling gauge invariance in the theory of the quantum Hall effect. III. The instanton vacuum and chiral-edge physics

A. M. M. Pruisken, B. Škorić, and M. A. Baranov*
Institute for Theoretical Physics, University of Amsterdam, Valckenierstraat 65, 1018 XE Amsterdam, The Netherlands
(Received 9 October 1998)

The concepts of an instanton vacuum and $\mathcal{F}$ invariance are used to derive a complete effective theory of massless edge excitations in the quantum Hall effect. Our theory includes the effects of disorder and Coulomb interactions, as well as the coupling to electromagnetic fields and statistical gauge fields. The results are obtained by studying the strong-coupling limit of a Finkelstein action, previously introduced for the purpose of unifying both integral and fractional quantum Hall regimes. We establish the fundamental relation between the *instanton vacuum* approach and the completely equivalent theory of *chiral edge bosons*. In this paper we limit the analysis to the integral regime. We show that our complete theory of edge dynamics can be used as an important tool to investigate long-standing problems such as long-range, smooth disorder, and Coulomb interaction effects. We introduce a two-dimensional network of chiral-edge states and tunneling centers (saddle points) as a model for smooth disorder. This network is then used to derive a mean-field theory of the conductances, and we work out the characteristic temperature (T) scale at which the transport crosses over from mean-field behavior at high T to the critical behavior plateau transitions at much lower T. The results explain the apparent lack of scaling which is usually seen in the transport data taken from arbitrary samples at finite T. Second, we address the problem of electron tunneling into the quantum Hall edge. We show that the tunneling density of states near the edge is affected by the combined effects of the Coulomb interactions and the smooth disorder in the bulk. We express the problem in terms of an effective Luttinger liquid with conductance parameter (g) equal to the filling fraction (ν) of the Landau band. Hence, even in the integral regime, our results for tunneling are completely non-Fermi-liquid-like, in sharp contrast to the predictions of single-edge theories. [S0163-1829(99)13739-1]

I. INTRODUCTION

In problems of quantum transport, symmetries play an important role. Recent advances in the theory of the quantum Hall effect primarily make use of electrodynamic gauge invariance as the fundamental symmetry of the strongly correlated electron gas.[1,2] This symmetry permits one to proceed with a minimum of microscopic input. Applications of Chern-Simons theory have been largely based upon phenomenological arguments. These applications have provided a universal language for the fractional quantum Hall effect in which the various hierarchy schemes could be treated on equal footing.[1]

Application of Chern-Simons theory has also led to the idea that many of the basic properties of incompressible quantum Hall states can be understood in terms of Luttinger liquid behavior of the edge excitations. This non-Fermi-liquid theory of edge excitations is now commonly used as a computational scheme for tunneling properties of different quantum Hall states as well as the thermodynamic properties of the fractionally charged quasiparticles. It is important to keep in mind, however, that unlike the conductance parameters, physical quantities like the tunneling density of states do not necessarily follow the rules of incompressibility. The lack of a microscopic theory of the fractional quantum Hall effect has led to controversial issues regarding the *definition* of the Hall conductance (notably for those states that have edge channels of opposite chirality).[3,4] Moreover, serious discrepancies have arisen between the predictions of the Luttinger-liquid theory of edge excitations[3] on the one hand, and experimental results on edge tunneling on the other.[5]

This paper is the third in a series in which we lay down the foundation for a microscopic theory of disordered compressible and incompressible states in the (fractional) quantum Hall regime. In previous papers[6,7] (hereafter called I and II) we introduced an effective Finkelstein action for localization and interaction effects. The Finkelstein action includes the topological concept of an instanton vacuum as well as the statistical (Chern-Simons) gauge fields. The inclusion of statistical gauge fields in the problem makes it possible to formulate a combined theory of composite fermions, localization and interaction effects. The results of *weak-coupling* analyses (both perturbative and nonperturbative, i.e. instantons) can then be used to obtain a global scaling diagram for the conductances. The integral as well as fractional quantum Hall regimes are incorporated in this scaling diagram. In this work, we are primarily interested in the strong-coupling limit of our action where the system has a gap in the density of states. This physical situation is the same as the one described by the Chern-Simons approach with one important exception: besides the Coulomb interactions, we also deal from first principles with the effects of disorder.

One of the main objectives of this work is to derive microscopically a Luttinger-liquid theory for edge excitations in the presence of disorder and electron-electron interactions. From our general, effective action point of view, we can say that the physics of edge excitations has a fundamental significance since it provides unique and invaluable information

 ©1999 The American Physical Society

on the topological concept of an ''instanton vacuum'' [8] in strong coupling.

An additional important advancement is that we obtain, for the first time, the *complete* Luttinger-liquid theory on the edge. We have the action for interacting chiral edge bosons coupled to external electromagnetic fields. This theory can now be used to define the Hall conductance in a general, unambiguous manner by expressing the appearance of an ''edge anomaly'' [9] in terms of Laughlin's gauge argument.[10]

The details of the analysis of edge excitations are described in Secs. III and IV. This analysis is based, to a large extent, on the various concepts which were introduced in I under the names ''$\mathcal{F}$ algebra'' and ''$\mathcal{F}$ invariance.'' Recall that in II we also studied these concepts, but in the weak-coupling regime. This paper therefore shows that $\mathcal{F}$ invariance retains its significance all the way down to the regime of strong coupling, where the massless excitations are confined to the edges of the sample. It is important to note that this symmetry is being demonstrated in the weak- as well as the strong-coupling regime.

The results of Secs. III and IV will serve as the starting point for a microscopic theory of edge excitations in the fractional quantum Hall effect. We shall limit ourselves here to the integer regime, since this already contains most of the difficulties. Extensions of our theory to include the fractional effect can be done by means of the statistical gauge fields. These will be reported elsewhere.

We shall begin by reviewing and extending the topological instanton vacuum approach to the quantum Hall effect, following the ordinary, free-electron replica formalism in Sec. II. In making the connection between topology and edge currents, we show that important aspects of the problem have previously been overlooked. In particular, we show that the *massless* excitations of the disordered edge states are obtained from *fluctuations* about *integer* quantized *topological charge* (Sec. II A 3). This important observation will serve as a starting point for most of the analyses in the remainder of this paper.

Massless edge excitations appear in the instanton vacuum theory for arbitrary number of field components (replicas) N_r and not just in the replica limit $N_r=0$. The present analysis revises our previously accumulated knowledge of the subject in at least two respects. First we recognize that a direct relationship exists between the numerical value of the instanton parameter θ (or σ_{xy}^0, Ref. 11) and the phenomenon of *interchannel scattering* at the edge. Here the number of edge channels equals the number of fully occupied Landau levels, and the phrase ''interchannel scattering'' refers to the effect of a random short-ranged potential.

Second, we review the earlier attempts to establish a general *topological principle* for quantization of the Hall conductance which includes the effect of localization of the bulk states. The mere existence of massless edge excitations turns out to have basic consequences for the *quantization phenomenon* which now can be shown to be a robust and fundamental aspect of the instanton vacuum theory with arbitrary values of N_r.

In all our work so far, we have substituted the phrase ''electronic disorder'' for a white noise random potential. This was always done for technical reasons alone. However, it is well known that in real quantum Hall devices slowly varying potentials are often present.[10,12] Till now these have in general been difficult to handle. Our microscopic theory of the edge enables us to treat long-range potentials as well as electron-electron interactions. In this paper we embark on solving two long-standing problems where smooth disorder and Coulomb interactions give rise to unexpected results. By addressing these problems we attack the core of the controversies that exist between the theory and experiments that presently span this subject.

The first problem we address is that of the plateau transitions. This we model as a percolating network of ''edge states'' (equipotential contours) and widely separated ''saddle points.'' A large class of such systems is then 'mapped' onto the nonlinear σ model representation for localization, and the main problem is to identify the length and energy scales of the ''bare'' parameters, or the mean-field conductances which together determine the renormalization starting point, i.e., the point where scaling occurs first. This starting point can involve, in principle, arbitrarily large distances and arbitrarily small energies, and this, obviously, complicates the observability of the critical behavior of the Anderson (plateau) transitions. We argue that Coulomb interaction effects lead to a modified mean-field theory of transport which is now observed in the experiments performed at finite temperatures. The chiral boson theory shall be used to actually compute the inelastic relaxation rate of the conducting electrons in the saddle-point network. This, then, might conceivably be the explanation for the empirical fits of the transport data taken recently from presently available samples.[13]

As the second typical example of long-ranged disorder effects, we embark on the problem of electron tunneling into the quantum Hall edge. We show that the Coulomb interactions between the edge and the ''localized'' bulk orbits dramatically differ from the predictions of theories which are based on isolated edges alone. Tunneling processes into the quantum Hall edge have, in fact, nothing to do with the quantization of the Hall conductance or the ''incompressibility'' statement which describe the nonequilibrium properties of the electron gas. We find that the tunneling density of states near the edge can be understood in terms of an effective edge theory which describes the equilibrium properties of the combined edge and bulk degrees of freedom. The Luttinger-liquid parameter g is related to the filling fraction ν of the bulk Landau level. This leads to a tunneling exponent which varies like $1/\nu$, in agreement with recent experimental data on the tunneling current, taken from samples in the fractional quantum Hall regime.[5] This situation is dramatically different from what is expected while assuming an isolated edge, or in the case of short-ranged disorder which gives rise to scattering between different edge states.[3]

In this paper and one that follows,[14] we carefully reexamine the consequences of interchannel edge scattering. We reproduce the completely different Kane-Fisher-Polchinsky[3] scenario of tunneling exponents in the integral and fractional regimes from our strong-coupling edge theory. However, we argue that both the assumptions (an isolated edge and short-ranged disorder or interchannel scattering) are clearly incorrect since the problem is two dimensional and dominated by long-ranged potential fluctuations as well as interaction effects.

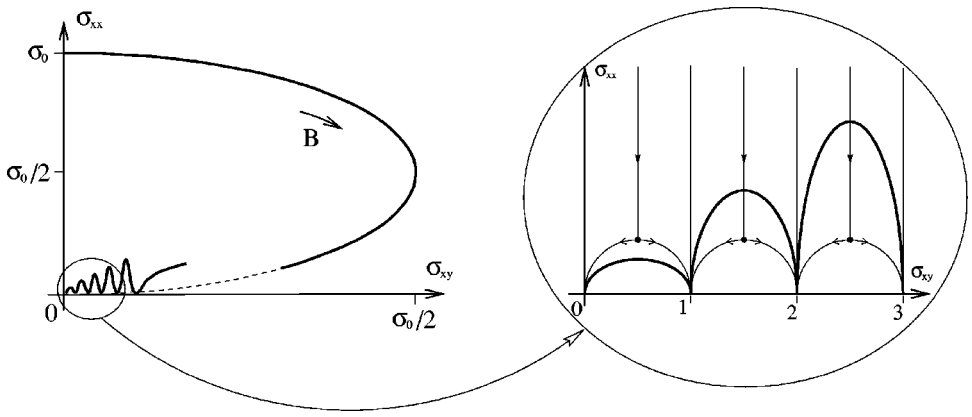

FIG. 1. Sketch of the mean-field conductances for a short-range disorder potential. The inset is the strong-field limit or quantum Hall regime. The renormalization group flow lines indicate how the mean-field theory results change after successive length scale transformations (after Refs. 8, 11, and 15).

The organization of this paper is as follows. In Sec. II we introduce the problem in the language of the replica free-electron theory. We briefly recall the instanton vacuum approach in Sec. II A 1.[8] The connection between topology and interchannel scattering between the chiral edge modes is made in Sec. II A 2. This leads to an exact solution of the instanton vacuum at the edge which can now be shown to be critical (Sec. II A 3).

In Sec. II B we introduce a two-dimensional network of chiral edge states as a model for the problem of long-ranged potential fluctuations. This is then used for mean-field purposes and for demonstrating universality of the plateau transitions in Sec. II B 1. In Sec. II B 2 we extend the network approach to include interaction effects. A semiclassical theory of transport is introduced in order to explain the lack of scaling recently found in many (ordinary) quantum Hall devices at finite temperatures. Section II B 3 contains several general remarks. In Sec. III we present a detailed derivation of the complete chiral edge theory using the fermionic path integral. In Section IV we make the fundamental connection between the *instanton vacuum* on the one hand, and the *Chern-Simons gauge theory* and *chiral edge bosons* on the other. In Sec. V we apply the theory of chiral edge bosons to several problems of long-range disorder and interaction in the bulk of the sample. These include the density of states for tunneling into the quantum Hall edge as well as the relaxation times entering into the transport problem of Sec. II B. We end this paper with a summary in Sec. VI.

II. EDGE EXCITATIONS

A. Sigma model

Let us recall the instanton vacuum theory[8,11] for the integral quantum Hall effect which is expressed in terms of the local-field variables $Q^{\alpha\beta}_{pp'}$, where $\alpha,\beta=1,\ldots,N_r$ are the replica indices and $p,p'=\pm 1$ are the indices denoting advanced or retarded waves. They can be represented as

$$Q=T^{-1}\Lambda T \quad \text{with} \quad \Lambda^{\alpha\beta}_{pp'}=\delta^{\alpha\beta}\delta_{pp'}\,\mathrm{sgn}(p) \tag{2.1}$$

and T a unitary matrix of size $2N_r\times 2N_r$. The complete action is given by

$$S[Q]=-\frac{1}{8}\sigma^0_{xx}\int d^2x\ \mathrm{tr}(\nabla Q)^2 + \frac{1}{8}\sigma^0_{xy}\int d^2x\ \mathrm{tr}\,\varepsilon_{ij}Q\partial_i Q\partial_j Q+\pi\rho_0\omega\int d^2x\ \mathrm{tr}\,\Lambda Q. \tag{2.2}$$

Here σ^0_{ij} stands for the mean field conductances in units of e^2/h (see Fig. 1), ρ_0 is the (exact) density of states at the Fermi energy and ω is the frequency. The second term in Eq. (2.2), proportional to the mean-field Hall conductance (σ^0_{xy}), has remained one of the most difficult chapters in the theory of Anderson localization in low dimensions. Most of the insight into the theory with $N_r=0$ number of field components has come from weak coupling renormalization theory (both perturbative and non-perturbative, i.e., instantons).[8] In particular we mention the global scaling diagram of the conductances as well as the appearance of a critical fixed point in the strong-coupling regime.[11] This fixed-point theory predicts a massless (metallic) phase at the Landau band center as well as the following scaling result for the conductances:[15]

$$\sigma_{ij}(L,B)=g_{ij}([L/\xi]^{1/\nu});\quad \xi=|B-B^*|^{-\nu} \tag{2.3}$$

which cannot be obtained in any different way. Here the function $g_{ij}(X)$ is a regular function of its argument, B^* is the critical magnetic-field strength, and ν stands for the critical index for the localization length ξ. Following the experimental tests of Eq. (2.3) by Wei *et al.*,[16] extensive numerical work on the free-electron gas has been performed, and the quoted best value for the critical index is $\nu=2.3$.[17]

To date, no exact (conformal) scheme for the critical indices exists. All that one can say at this time is that the field of exactly solvable models is not sufficiently developed to be able to handle the specific subtleties of topology and replica field theory. These subtleties are all well understood within the elaborate framework of weak coupling expansion techniques,[8] and the results were used to unfold and predict the entire singularity structure of the theory, notably Eq. (2.3).

In previous work[18] we have shown that the theories of free and interacting electrons share the same basic features such as asymptotic freedom, instantons, etc. The same scaling diagram for the conductances was obtained, which means that Eq. (2.3) also remains valid when the Coulomb interactions are taken into account. This important result was con-

jectured but otherwise not understood at the time of the original experiments on criticality.

1. Strong coupling

In this paper we address the subtleties of the instanton vacuum theory in an extremely important exactly solvable limit where $\rho_0=\sigma^0_{xx}=0$ and where the Hall conductance is integer quantized ($\sigma_{xy}=m$). Physically this happens when the Fermi energy is located in a density of states gap between adjacent Landau bands. In this *strong-coupling* limit massless excitations do exist at the edges of the system. Since several basic aspects of the problem have previously gone unnoticed, we shall proceed first within the free-electron formalism of Eqs. (2.1) and (2.2). We come back to the fermionic path integral in Secs. III and IV.

For m completely filled Landau levels the action becomes simply

$$S[Q]=\frac{m}{8}\int d^2x\ \mathrm{tr}\,\varepsilon_{ij}Q\partial_i Q\partial_j Q=\frac{m}{2}\oint d\vec{x}\cdot\mathrm{tr}(\Lambda T\nabla T^{-1}), \tag{2.4}$$

where the surface integral is taken over the sample's edge. Recall that Eq. (2.4) is quantized according to

$$S[Q]=2\pi i m q[Q], \tag{2.5}$$

with q the integer topological charge, provided that the T matrix reduces to a $\mathrm{U}(N_r)\times\mathrm{U}(N_r)$ gauge at the edge.[11]

Under these circumstances the sample edge has been contracted to a single point (spherical boundary conditions) and Eq. (2.5) is a realization of the formal homotopy theory result $\pi_2(G/H)=Z$ which states that the mapping of Q onto the two dimensional (2D) plane is described by a set of integers q. It is natural to take the theory one step further and propose the quantization of the charge $q[Q]$ as the topological principle in replica field theory which forces the Hall conductance (m) itself to be integer quantized. The idea has led to a consistent quantum theory of conductances that unifies a fundamental aspect of asymptotically free-field theory (i.e., dynamic mass generation) with the quantum Hall effect.[8] More specifically, it says that the conductances in Eq. (2.3) always scale toward $\sigma_{xx}=0$, $\sigma_{xy}=m$ for L large enough.

One can show[11] that the $\mathrm{U}(N_r)\times\mathrm{U}(N_r)$ gauge condition at the edge is the replica field theory version of a static U(1) gauge acting on the physical edge states. Such a U(1) gauge implies that an integer number of edge levels has crossed the Fermi level. This level crossing is necessarily induced by the averaging procedure over random potentials.

Nevertheless, it is somewhat disappointing to know that the topological invariant in Eq. (2.2), as it was discovered originally in a microscopic derivation, is truly defined with *free* boundary conditions and without any separation between edge and bulk degrees of freedom.[8] So far, the precise significance of boundary conditions has remained obscure.

2. Interchannel edge scattering

In what follows, we show that the fluctuations about precisely quantized values for the topological charge represent, in fact, essential physics of the problem, since they describe the dynamics of (massless) edge excitations. In order to see this, we write T as the product of a $\mathrm{U}(N_r)\times\mathrm{U}(N_r)$ gauge U, and a small fluctuation t:

$$T=Ut. \tag{2.6}$$

The action now becomes

$$S[Q]=2\pi i m q[U]+\frac{m}{2}\oint d\vec{x}\cdot\mathrm{tr}(\Lambda t\nabla t^{-1})$$
$$+\pi\rho_{\mathrm{edge}}\omega\oint dx\ \mathrm{tr}\,\Lambda Q, \tag{2.7}$$

with ρ_{edge} the density of edge states. One way of identifying Eq. (2.7) as the effective theory of disordered chiral edge states is to redo the derivation, but now for the 1D system with Hamiltonian

$$\mathcal{H}_{\mathrm{edge}}=-iv_d\partial_x+V(x), \tag{2.8}$$

where v_{d} is the drift velocity of the edge electrons and $V(x)$ the random potential. It turns out that our initial guess (2.8) is correct only in the case $m=1$ in Eq. (2.7). This problem is easily resolved once one realizes that m really stands for the number of filled Landau levels, such that Eq. (2.8) should be replaced by a Hamiltonian for a total of m edge channels. Hence, an obvious second guess would be

$$\mathcal{H}_{\mathrm{edge}}=\sum_{j=1}^{m}\mathcal{H}^{(j)}_{\mathrm{edge}}, \tag{2.9}$$

where $\mathcal{H}^{(j)}_{\mathrm{edge}}$ is the same for all j, i.e., each of the m eigenstates experiences the same white noise potential $V(x)$, just as it appears in the original problem in two spatial dimensions. This, however, is not correct and the theory with general m, [Eq. (2.7)], necessarily requires interchannel scattering to take place. We have to start from a matrix Hamiltonian

$$\mathcal{H}^{jj'}_{\mathrm{edge}}=-iv_d\delta_{jj'}\partial_x+V_{jj'}(x), \tag{2.10}$$

where V is a Hermitian matrix. The matrix elements $V_{jj'}$ connect the edge channels j and j' and are distributed with a weight

$$P[V]\propto\exp\left\{-\frac{1}{g}\oint dx\ \mathrm{tr}\,V^2\right\}. \tag{2.11}$$

One can construct a generating function for the free particle Green's functions as usual, according to

$$Z=\int\mathcal{D}[\bar{\psi}\psi]\int\mathcal{D}[V]P[V]\exp\sum_{p=\pm,\alpha,jj'}\oint dx\ \bar{\psi}^{\alpha,j}_p$$
$$\times[(\mu+ip\omega)\delta_{jj'}-\mathcal{H}^{jj'}_{\mathrm{edge}}]\psi^{\alpha,j'}_p. \tag{2.12}$$

In Appendix C we show that Eqs. (2.12) and (2.7) are identical in the limit of large distances.

3. Criticality at the edge

We next point out that the results of the previous section provide an exact solution to our topological theory at the edge [Eqs. (2.4)–(2.7)] for all values of N_r. The simple but

important observation to be made is that the random potential $V_{jj'}(x)$ in Eq. (2.12) can be ''gauged away,'' i.e. absorbed in a redefinition of the fermion fields, and all that remains is the trivial theory of ''pure'' chiral edge states,

$$Z=\int \mathcal{D}[\bar{\psi}\psi]\exp\sum_{p=\pm,\alpha,j}\oint dx\ \bar{\psi}_p^{\alpha,j}[-iv_d\partial_x+i\omega p]\psi_p^{\alpha,j}. \tag{2.13}$$

Equation (2.13) is just a formal way of saying that edge electrons do not Anderson localize, because chirality excludes backscattering processes on random impurities. Following up on the analysis of Appendix C we will use the simplicity of Eq. (2.13) and derive explicit results for the Q field (2.7). Write

$$G_{\pm}^{jj'}(x,x')=\langle x,j|(-\mathcal{H}_{\rm edge}\pm i\omega)^{-1}|x',j'\rangle, \tag{2.14}$$

$$G_{\pm}(x,x')=\langle x|(iv_d\partial_x\pm i\omega)^{-1}|x'\rangle \tag{2.15}$$

to represent the single-particle propagator of the dirty edge [Eq. (2.12)] and clean edge [Eq. (2.13)], respectively. Some useful identities are given by

$$\rho_{\rm edge}=\frac{1}{2\pi i}\sum_j\ [G_-^{jj}(x,x)-G_+^{jj}(x,x)]$$
$$=\frac{m}{2\pi i}[G_-(x,x)-G_+(x,x)]. \tag{2.16}$$

Here $\rho_{\rm edge}$ denotes the density of edge states at the Fermi level which can be obtained explicitly from the right-hand side,

$$\rho_{\rm edge}(x)=\rho_{\rm edge}=m/2\pi v_d. \tag{2.17}$$

Equation (2.17) shows that the density of edge electrons is a constant, independent of x and disorder, as it should be. An important conclusion now follows for the theory of Q fields [Eq. (2.7)], namely,

$$\langle Q\rangle=\Lambda \tag{2.18}$$

[where the expectation is with respect to Eq. (2.7)], which holds for arbitrary N_r. This result may be obtained e.g., by differentiating both theories [Eq. (2.7) and Eqs. (2.12), (2.13)] with respect to ω. Notice that Eq. (2.18) can be regarded as the ''order parameter'' (analogous to the magnetization in the language of the Heisenberg ferromagnet), and one would naively expect this quantity to vanish in one spatial dimension. The result $\langle Q\rangle=\Lambda$ indicates, however, that the continuous symmetry is permanently broken at the edge of the instanton vacuum for all numbers of field components N_r. This apparent violation of the Mermin-Wagner-Coleman theorem is clearly due to the lack of positive definite Boltzmann weights in our problem that is described by an imaginary action [Eq. (2.7)]. Equation (2.18) also indicates that the edge of the topological vacuum is critical. The simplest way of demonstrating this is by employing the background field method. For example, the replacement $t\to tt_0$ in the second term of Eq. (2.7) can be written as

$$\oint dx\ {\rm tr}[\Lambda t\partial_x t^{-1}]\to\oint dx\ {\rm tr}[\Lambda tt_0\partial_x(t_0^{-1}t^{-1})]$$
$$=\oint dx\ {\rm tr}[\Lambda t\partial_x t^{-1}]$$
$$+\oint dx\ {\rm tr}[Qt_0\partial_x t_0^{-1}]. \tag{2.19}$$

Here, t_0 represents a fixed and slowly varying background field. We obtain an effective action for t_0 as follows:

$$S_{\rm eff}[t_0]=\frac{m}{2}\oint dx\ {\rm tr}[t_0\partial_x t_0^{-1}\langle Q\rangle]$$
$$=\frac{m'}{2}\oint dx\ {\rm tr}[\Lambda t_0\partial_x t_0^{-1}]. \tag{2.20}$$

Equation (2.20) defines an ''effective'' parameter $m'=m\ {\rm tr}\,\Lambda\langle Q\rangle/2N_r$ which can be identified as the ''Hall conductance'' and which provides information on the renormalization of the theory at large distances.[19] Apparently we have $m'=m$. The same conclusion can be drawn for the ω parameter (i.e., $\omega'=\omega\ {\rm tr}\,\Lambda\langle Q\rangle/2N_r=\omega$), and hence we are dealing with a critical fixed point theory. The full significance of this result will become clear in the forthcoming sections, where we make contact with the theory of chiral edge bosons.

For the remainder of this section we will elaborate on several other identities and relations that will be used later on. The most important pair correlation of the Q fields can be obtained as follows:

$$N(x,x')=\pi^2\rho_{\rm edge}^2\langle Q_{+-}^{\alpha\beta}(x)Q_{-+}^{\beta\alpha}(x')\rangle$$
$$=\sum_{jj'}\ G_-^{jj'}(x,x')G_+^{j'j}(x',x)$$
$$=mG_-(x,x')G_+(x',x). \tag{2.21}$$

Here α and β are fixed but arbitrary replica channels, and

$$G_-(x,x')G_+(x',x)=\frac{i}{v_{\rm d}}\int\frac{dk}{2\pi}\frac{e^{ik(x'-x)}}{v_dk+2i\omega}$$
$$=\frac{1}{v_d^2}\theta(x'-x)$$
$$\times\exp\left[-\frac{2\omega}{v_d}(x'-x)\right]. \tag{2.22}$$

The step function θ shows that a chiral electron, being created at position x and drifting in the positive direction, can only be destroyed at a ''later'' position $x'>x$. Notice that we have the standard sum rule

$$\int dx'N(x,x')=\pi\rho_{\rm edge}/\omega. \tag{2.23}$$

The other pair correlations of the Q-fields vanish identically. In particular, it is straightforward to show that

$$\langle Q_{pp}^{\alpha\beta}(x)Q_{p'p'}^{\gamma\delta}(x')\rangle_{\rm cum}=0 \tag{2.24}$$

for all $p,p'=\pm$ and all replica channels α,β,γ, and δ. Next we wish to clarify the significance of several Q-field operators that have appeared in different contexts before. First, there are the higher order corrections to the theory of Eq. (2.7) of the type (see Appendix C)

$$\mathrm{tr}\left[\frac{m}{2}\partial_x+\pi\omega\rho_{\mathrm{edge}}\Lambda,Q\right]^2. \tag{2.25}$$

Second, we mention the bilinear combinations of the form

$$A_1\,\mathrm{tr}\,\Lambda Q\,\mathrm{tr}\,\Lambda Q+A_2\,\mathrm{tr}[\Lambda,Q][\Lambda,Q], \tag{2.26}$$

which are known to describe the anomalous fluctuations in the density at the quantum Hall transitions, as well as in the localization problem in $2+\varepsilon$ dimensions.[20] We have already seen, however, that the density of chiral electrons does not fluctuate as one moves along the edge and we therefore expect Eq. (2.26) to be irrelevant. A classification of these operators follows from the classical equations of motion of the topological action (2.7), which can be written as

$$\left[\frac{m}{2}\partial_x+\pi\omega\rho_{\mathrm{edge}}\Lambda,Q\right]=0. \tag{2.27}$$

This immediately implies that the higher-dimensional operators [Eq. (2.25)] are, in fact, *redundant*. Next, from the identity

$$\int^x dx'\,\mathrm{tr}\,\Lambda Q(x')\mathrm{tr}\left[\frac{m}{2}\partial_{x'}+\pi\omega\rho_{\mathrm{edge}}\Lambda,Q(x')\right]\Lambda=0, \tag{2.28}$$

it directly follows that the first term in Eq. (2.26) is redundant as well. Finally, from Eq. (2.27) one also obtains

$$\int^x dx'\,\mathrm{tr}[\Lambda,Q(x')]\left[\frac{m}{2}\partial_{x'}+\pi\omega\rho_{\mathrm{edge}}\Lambda,[\Lambda,Q(x')]\right]=0, \tag{2.29}$$

and it is readily seen that the second operator in Eq. (2.26) is also redundant.

B. Plateau transitions revisited

1. Long-range potential fluctuations

In this section we show how the notion of critical edge states can be used in order to gain insight into the problem of ''long-ranged potential fluctuations.'' This long-standing problem, which is very difficult to handle within the formal nonlinear σ-model methodology, plays an extremely important role experimentally. For instance, it has been stressed many times and at many places elsewhere that the plateau transitions as observed in the detailed experiments of Wei *et al.*[16] are very difficult to observe in general in arbitrary samples, due to the presence of slowly varying potential fluctuations.

A slowly varying potential is the generic type of disorder in the standard GaAs heterostructure, which has historically led to semiclassical considerations (percolation picture) of delocalization near the Landau band center.[21] It is important to recognize that also our critical system [Eq. (2.3)] is very sensitive to the presence of smooth potentials (or ''inhomogeneities'') in the sample. For example, the critical magnetic field B^* may be slowly varying throughout the system due to inhomogeneities in the electron density. This means that the scaling result is valid only up to a certain fixed value for L. Beyond this value the remaining ''extended'' states in the problem may be confined to the equipotential contours of the inhomogeneity potential, quite similar to the semiclassical picture of percolation.

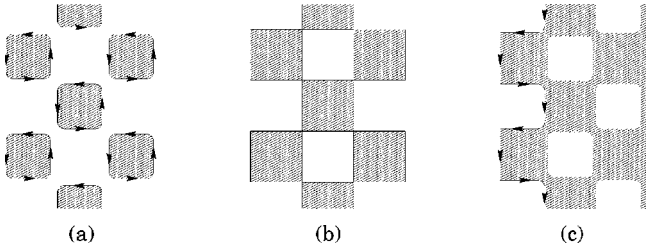

FIG. 2. Backbone cluster as a network of saddle points. Shaded areas have $\nu=1$, white areas $\nu=0$. The arrows indicate the direction of the currents. (a) Less than half-filling. (b) Exactly half-filling. (c) A filling fraction larger than one-half.

It is generally difficult to obtain detailed knowledge on the various length and energy scales that are involved in the crossover problem between percolation and localization. In what follows, we present the simplest possible scenario for crossover that enables us to deal simultaneously with interaction effects and such basic concepts as ''mean-field theory'' and ''universality'' of the plateau transition.

2. Quantum percolation

In order to fix the thought, we imagine the equipotential contours near half filling to form a large cluster (Fig. 2).[22] Since the disconnected, closed contours do not contribute to the transport, we focus our attention to an infinite backbone cluster which we take as a regular 2D array of saddle points, and we disregard all the loose hanging, finite pieces [Fig. 2(b)]. The saddle points (the sites of the square lattice) are connected to one another by the disordered 1D chiral edge channels (links on the lattice). This network can alternatively be looked upon as a checkerboard with filling fractions alternating between the values $\nu=0$ and $\nu=1$. The kinetic part of the action for this system may be written in the form of Eq. (2.4),

$$S[Q]=\frac{1}{8}\int d^2x\,m(\vec{x})\,\mathrm{tr}\,\varepsilon_{ij}Q\partial_iQ\partial_jQ, \tag{2.30}$$

with $m(\vec{x})=0,1$ [Fig. 2(b)]. Using the parametrization of Eq. (2.6) the action can also be written in the form of Eq. (2.7) which is now solely defined on the links of the square lattice:

$$S[Q]=2\pi iq[U]+\frac{1}{2}\sum_i\oint_i dx\,\mathrm{tr}(\Lambda t\partial_x t^{-1})$$
$$+\pi\omega\rho_{\mathrm{link}}\sum_i\oint_i dx\,\mathrm{tr}\,\Lambda Q. \tag{2.31}$$

Here the sum is over all the black squares, and the integrals are over the contours of the black squares. Despite the fact that this action does not contain any dissipative (σ_{xx}) terms,

it is easy enough to show that in the long-wavelength limit, Eq. (2.30) reduces to the form of the sigma model action (2.2), with

$$\sigma_{xx}^0 = 1/2, \quad \sigma_{xy}^0 = 1/2. \tag{2.32}$$

The reason for this is contained in the fact that the saddle points act like scattering centers which render the system dissipative at large distances. In order to demonstrate this, all one needs to do is to follow up on Eq. (2.20), where the background field t_0 now represents the ''slow modes'' that are kept. The t field variables are the ''fast modes'' which contain all the wavelengths smaller than the lattice constant, i.e., the average distance between the saddle points, and which are integrated out. This leads to an effective action for each link according to

$$\begin{aligned} S_{\text{link}}[t_0] &= \frac{1}{2}\int_{\text{link}} dx\ \text{tr}(\langle Q\rangle t_0 \partial_x t_0^{-1}) \\ &\quad + \frac{1}{8}\left\langle \left[\int_{\text{link}} dx\ \text{tr}(Q t_0 \partial_x t_0^{-1}) \right]^2 \right\rangle_{\text{cum}} \\ &= \frac{1}{2}\int_{\text{link}} dx\ \text{tr}(\Lambda t_0 \partial_x t_0^{-1}) - \frac{\sigma_{xx}^0}{8}\int_{\text{link}} dx\ \text{tr}(\partial_x Q_0)^2, \end{aligned} \tag{2.33}$$

where $Q_0 = t_0^{-1}\Lambda t_0$ and the expectation is with respect to the theory [Eq. (2.7)] with $m=1$. The subscript ''cum'' indicates that only connected diagrams are taken. The $\sigma_{xx}^0 = L_0/2$ is the 1D conductivity of a single channel of length L_0, a well-known result in the theory of pure metals. These results are obtained by making use of Eq. (2.18) as well as Eqs. (2.21)–(2.24) in the limit $\omega = 0$. Next, by taking the sum over all links one can absorb the factor L_0 into the definition of a 2D integral,

$$-\frac{1}{16}\sum_{\text{links}} L_0 \int_{\text{link}} dx\ \text{tr}(\partial_x Q_0)^2 \to -\frac{1}{16}\int d^2x\ \text{tr}(\nabla Q_0)^2. \tag{2.34}$$

Here we only used the fact that the Q_0 field variable varies slowly over a distance L_0. The first term in Eq. (2.33) can be handled in a similar way. For instance, it can be rewritten in the form of Eq. (2.30) with Q replaced by Q_0, which is then followed by taking the continuum limit according to

$$\begin{aligned} \frac{1}{2}\sum_{\text{links}} \int_{\text{link}} dx\ \text{tr}(\Lambda t_0 \partial_x t_0^{-1}) &\to \frac{1}{8}\int d^2x\ m(\vec{x})\varepsilon_{ij}\ \text{tr}\, Q_0 \partial_i Q_0 \partial_j Q_0 \\ &\to \frac{1}{16}\int d^2x\, \varepsilon_{ij}\ \text{tr}\, Q_0 \partial_i Q_0 \partial_j Q_0. \end{aligned} \tag{2.35}$$

The result of Eqs. (2.33)–(2.35) is identical to the statement made in Eq. (2.32). Notice that Eq. (2.32) is precisely the point where we expect the σ model action (2.2) in the limit $N_r = 0$ to have a critical phase. Hence, we have established a direct connection between critical 1D edge states on the one hand and the 2D delocalization transition of the band center on the other. It is important to stress that this connection has the following ingredients.

(1) The infinite percolation cluster at the band center contains a finite density of saddle points. This translates into a finite density of scattering centers which, in turn, is responsible for making the sample diffusive (dissipative) at large distances.

(2) The parameters σ_{xx}^0 and σ_{xy}^0 [Eq. (2.32)] constitute a mean-field theory of the conductances which is valid for length scales L_0. This holds for any value of N_r and not just for $N_r = 0$.

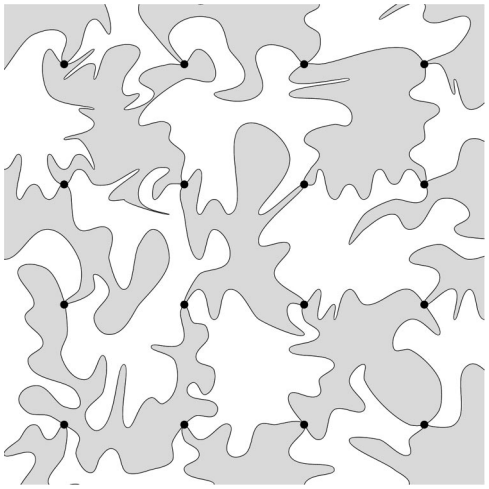

FIG. 3. Backbone cluster as in Fig. 2(b), but with highly ramified contours between saddle points (●).

Without going into further detail we mention the fact that the analysis can easily be generalized to more complicated situations. For example, the links between the saddle points need not be straight lines. They can be taken as arbitrarily complex, non-intersecting paths reflecting the highly ramified percolation contours (Fig. 3). The same result [Eq. (2.32)] applies to all cases, indicating that the general result $\sigma_{xx}^0 = 1/2$ actually stands for the quantized conductance in one dimension.

3. Mean-field theory

Next we wish to extend our mean-field analysis [Eq. (2.32)] to include also the energies away from the Landau

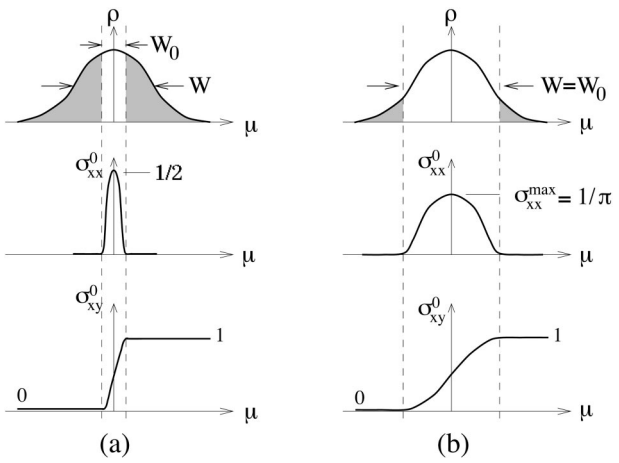

FIG. 4. Mean-field theory for the lowest Landau level, with varying chemical potential μ. (a) Smooth long-range disorder. (b) Short-range disorder (see text).

band center. For this purpose we have to relate the range in energy W_0 within which the equipotential contours form an infinite saddle-point cluster to the total bandwidth W of the Landau band. It is understood that the phrase ''saddle point'' actually stands for those special points where two equipotential contours approach each other at a distance of the order of the magnetic length l_0 or smaller. By assuming a simple quadratic form for the potential near saddle points, we obtain the estimate

$$W_0 \approx (l_0/\lambda)^2 W, \tag{2.36}$$

where λ is the characteristic correlation length of the random potential, which we have taken to be much larger than l_0, and W equals the amplitude of the potential fluctuations. The σ-model theory or, equivalently, the scaling theory of localization, only applies to the (narrow) energy band W_0 about the band center. For energies just outside W_0 the network of saddle points is broken up into disconnected islands of size $L_0 \times L_0$ [Figs. 2(a) and 2(c)]. The absence of any quantum tunneling means that no correlation exists between the islands (they are represented by independent actions as long as one works within the free-electron approach). In the language of the σ model, the situation is represented by putting $\sigma_{xx}=0$ but $\sigma_{xy}=m=$ integer. The latter follows from the long-ranged correlations which still exist near the edge and which can generally be expressed in terms of an integer number m of edge channels. In Fig. 4(a) we illustrate the behavior of the density of states ρ and the conductances σ_{ij}^0 as a function of energy μ at zero temperature.

The σ-model conductance parameters σ_{ij}^0 can be expressed as a function of the dimensionless quantity $\Delta\mu/W_0$,

$$\sigma_{ij}^0 = f_{ij}(\Delta\mu/W_0), \tag{2.37}$$

where $\Delta\mu$ is the energy relative to the Landau band center. The f_{ij}'s are nonuniversal and generally depend on the microscopic details of the randomness. For comparison we have plotted the results of the more familiar theory of short-ranged scatterers (self-consistent Born approximation) in Fig. 4(b). In this case, there is only a small difference between W_0 and W due to the localized states in the Gaussian tails of the Landau band.

An estimate for L_0 can be obtained as follows. Let $|\Delta\mu| \approx W_0$ denote the energies where the saddle point breaks up into disconnected equipotential contours of size $L_0 \times L_0$ [Figs. 2(a) and 2(c)]. According to the semiclassical picture of percolation we can relate the typical cluster size ξ_p to the energy $\Delta\mu$ according to

$$\xi_p \sim \lambda(\Delta\mu/W)^{-4/3}, \tag{2.38}$$

where the critical index 4/3 is the exponent for semiclassical localization. By identifying the points $|\Delta\mu|=W_0$ and $\xi_p = L_0$ in Eq. (2.38) we obtain the estimate

$$L_0 \approx l_0(\lambda/l_0)^{11/3} \quad (\lambda \gg l_0), \tag{2.39}$$

or, more generally,

$$W_n \approx \frac{l_n}{\lambda} W, \quad L_n \approx l_n(\lambda/l_n)^{11/3} \quad (\lambda \gg l_n). \tag{2.40}$$

The λ is an adjustable parameter in the theory, and it ranges between microscopic distances ($l_0 \approx 100$ Å) and infinity.

4. Interaction effects

It is quite possible that L_0 [Eq. (2.39)] is many times larger than the micron regime which is the typical scale for inelastic processes at low temperatures. This means that the critical behavior [Eq. (2.3)] cannot be observed within the limitations of ordinary laboratory experiments. This, then, is the easiest and crudest explanation for the lack of scaling in many samples. As a first step toward a more quantitative understanding of transport at finite T, we come back to the distinction, made in the beginning, between the backbone cluster and the disconnected, ''loose hanging'' pieces. Due to the electron-electron interactions, motion of the conducting electrons on the saddle-point network is affected by the localized electrons. This may be expressed in terms of a *relaxation time* $\tau_{\rm in}$ which is a characteristic time for equilibration between the conducting and localized electrons. Later on in this paper (Sec. V D) we shall address the problem of interaction effects and show that

$$1/\tau_{\rm in} = \beta_1 T + \beta_2 T^2 + \cdots \tag{2.41}$$

at low temperatures. This expression is determined by the collection of ''nearly saddle points'' where quantum tunneling is not possible but where the interactions between the conducting and localized particles are strongest nevertheless. The importance of ''nearly saddlepoints'' can be seen by comparing the wave functions at different energies close to the Landau band center. What is a saddle-point configuration at one energy may turn into a ''nearly saddle point'' at another, and vice versa. These abrupt changes in the configuration of the conducting network at slightly different energies blur the distinction between saddlepoints and ''nearly saddle point'' configurations as far as finite temperatures are concerned. This means that the relaxation time $\tau_{\rm in}$ [Eq. (2.41)] determines an effective bandwidth $W_{\rm eff}=W_0+\tau_{\rm in}^{-1}$ of states that contribute to the conduction at finite temperatures. Equation (2.37) is replaced by the expression

$$\sigma_{ij}^0(T)=f_{ij}(\Delta\mu/W_{\rm eff})=f_{ij}(\Delta\mu/[W_0+\tau_{\rm in}^{-1}]). \tag{2.42}$$

This result is a characteristic feature of long-ranged potential fluctuations, and it does not occur in the problem of short-ranged scatterers. To conclude this section, we shall next estimate the range of validity of result (2.42). Write

$$v_d\tau_{\rm in}=L_{\rm in}, \quad v_d\approx 2\pi l_0^2 W/\lambda. \tag{2.43}$$

$L_{\rm in}$ is the mean free path for drifting along the links of the lattice. We mentioned earlier that the actual path between two saddle points is arbitrarily convoluted and very long. Let L_t denote the actual path length between saddle points; then the criterion for scaling is clearly given by

$$L_{\rm in}>L_t. \tag{2.44}$$

Next we use the ramification hypothesis[10] in order to relate L_t to the shortest distance between saddle points (L_0). We obtain

$$L_t\propto L_0^{\sigma}, \tag{2.45}$$

with σ somewhere between 1 and 2. The criterion for scaling [Eq. (2.44)] now implies

$$\tau_{\rm in}^{-1}<(l_0/\lambda)^{8\sigma/3}W_0\ll W_0. \tag{2.46}$$

This result indicates that Eq. (2.42) is very likely to be observed in the (many) samples that are characterized by a smooth disorder potential. The results of this section are consistent with the recently reported empirical fitting[13] of the transport data in the quantum Hall regime. Since we are necessarily operating with an almost complete lack of knowledge on the microscopic details of sample disorder, it is conceivable that other types of inhomogeneity, especially those in low-mobility samples, explain the same thing.

5. Modified σ-model representation

The subjects of critical edge states as well as long-ranged disorder have left several conceptual questions that still need to be answered. For example, we have seen that short-ranged disorder causes interchannel scattering between the chiral edge states. Since we do not expect interchannel scattering to occur when the potential fluctuations are smooth (relative to the magnetic length), it is necessary to reinvestigate the meaning of instanton vacuum theory [Eq. (2.2)] for $\nu>1$ ($\sigma_{xy}^0>1$). Scattering between multiple edge states is avoided by writing, instead of Eq. (2.2),

$$S_{\rm eff}[Q^{(n)}]=\sum_{n=0}^{\infty}\left[-\frac{1}{8}\sigma_{xx}^{(n)}\int d^2x\,{\rm tr}[\nabla Q^{(n)}]^2+\frac{1}{8}\sigma_{xy}^{(n)}\int d^2x\,{\rm tr}\,\varepsilon_{ij}Q^{(n)}\partial_i Q^{(n)}\partial_j Q^{(n)}+\pi\rho^{(n)}\omega\int d^2x\,{\rm tr}\,\Lambda Q^{(n)}\right], \tag{2.47}$$

where the sum runs over all the Landau levels n. The $Q^{(n)}$ stands for an independent-field variable Q for each Landau level separately. The $\sigma_{ij}^{(n)}$'s are the nth Landau-level contributions to the mean-field conductances, which are now given by

$$\sigma_{ij}=\sum_{n=0}^{\infty}\sigma_{ij}^{(n)}. \tag{2.48}$$

The $\sigma_{ij}^{(n)}$'s are all the same [Fig. 4(a)] except for an appropriate shift in energy. Since $0\leqslant\sigma_{xy}^{(n)}\leqslant 1$ for each n, it is clear that Eq. (2.47) is the appropriate generalization of the theory (Sec. II A) to include filling fractions larger than one. The theories of Eqs. (2.47) and (2.2) are identical as far as the critical behavior of the plateau transitions is concerned. Equation (2.47) cannot, however, be used in the limit of small magnetic field, where the Landau levels partly or completely overlap. The details of crossover require a separate analysis.

6. Topological principle

In Refs. 11 and 23 a *topological principle* for Hall quantization was introduced. The basic idea is to relate the concept of dynamic mass generation in asymptotically free field theories to the quantization of the Hall conductance, which is now recognized as a universal quantum phenomenon at macroscopic length scales. The formulation presented in Refs. 11 and 23 is actually incomplete because the subtleties of edge effects were not sufficiently understood at that time. In order to see whether the instanton vacuum approach is, in fact, free of ambiguities, we shall follow up on the background field method which is known to generate the Kubo formulas for the conductances. Write

$$\exp S_{\rm eff}[t_0]=\int\mathcal{D}[Q]\exp(S_0[t_0^{-1}Qt_0]+\pi\rho_0\omega\,{\rm Tr}\,\Lambda Q), \tag{2.49}$$

where

$$S_0[Q]=-\frac{1}{8}\sigma_{xx}^0\,{\rm Tr}(\nabla Q)^2+\frac{1}{8}\sigma_{xy}^0\,{\rm Tr}\,\varepsilon_{ij}Q\partial_i Q\partial_j Q. \tag{2.50}$$

Equation (2.49) defines an effective action $S_{\rm eff}$ for the fixed and slowly varying matrix field t_0. One can show that $S_{\rm eff}$ is of the same form as S_0, i.e.,

$$S_{\rm eff}[t_0]=-\frac{1}{8}\sigma_{xx}\,{\rm Tr}(\nabla Q_0)^2+\frac{1}{8}\sigma_{xy}\,{\rm Tr}\,\varepsilon_{ij}Q_0\partial_i Q_0\partial_j Q_0, \tag{2.51}$$

with $Q_0=t_0^{-1}\Lambda t_0$. Equation (2.51) is actually the only possible action that respects the global $U(2N_r)$ symmetry as well as the local $U(N_r)\times U(N_r)$ gauge invariance of the problem. The main problem next is to obtain explicit knowledge of the ''effective'' parameters σ_{ij} in (Eq. 2.51) which now represent the (exact) Kubo expressions for the conductances. As long as one works with spherical boundary conditions on the matrix field Q (which have been assumed from the start), the quantization of the Hall conductance is readily established. All that one needs in fact is that the theory develops a mass gap in the limit of large distances. The insertion of slowly varying background fields (with $Q=\Lambda$ at the

edge) should then leave the theory unchanged in the limit $\omega \to 0$. This, then, directly leads to the statement that $\sigma_{xx} = 0$ and $\sigma_{xy} =$ integer.

The renormalization-group flows, obtained from instanton calculations, can next be used to show how the conditions of the quantum Hall effect appear as stable, infrared fixed points of the theory for arbitrary number of field components N_r. Although spherical boundary conditions are naturally imposed on the weak-coupling problem due to the finite action requirement of topological excitations, they are, however, controversial in the strong-coupling regime.

Armed with the insight gained from edge excitations in the previous sections, we next apply the background field procedure to the theory, but now with free-boundary conditions on Q, as it should be. For the special case where the Fermi energy lies in a density of states gap, Eq. (2.49) has already been addressed in Sec. B 2. $S_{\rm eff}$ for arbitrary N_r is given by

$$S_{\rm eff}[t_0] = 2\pi i m q[Q_0] - \frac{m^2}{32\pi\omega\rho_{\rm edge}} \oint dx\, {\rm tr}(\partial_x Q_0)^2, \tag{2.52}$$

where $q[Q_0] = 1/16\pi i\, {\rm Tr}\, \varepsilon_{ij} Q_0 \partial_i Q_0 \partial_j Q_0$, and the contour integral is along the sample edge. Comparing Eqs. (2.52) and (2.51), we see that the quantum Hall conditions are satisfied, but there are additional edge terms which are clearly the result of the chiral edge modes in the problem. Equation (2.52), in the limit $\omega \to 0$, forces the background field to obey the classical equations of motion (defined along the sample edge)

$$\partial_x Q_0 = 0. \tag{2.53}$$

The solution $Q_0 =$ constant at the edge simply means that spherical boundary conditions are automatically enforced by the chiral edge excitations. Notice that the effect of $S_{\rm eff}$ reduces to that of a phase factor which is immaterial provided the Hall conductance m precisely equals an integer. Physically, this phase factor arises from an integer number of edge electrons that have crossed the Fermi level as a result of the background field insertion.[11]

The same procedure can be repeated for the theory with $\sigma^0_{xx} \neq 0$, making use of the fact that a mass gap exists in the system of long-wavelength excitations, i.e. a finite localization length ξ. One expects Eq. (2.52) to be modified according to

$$S_{\rm eff}[t_0] = -\frac{\sigma_{xx}}{8} {\rm Tr}(\nabla Q_0)^2 + 2\pi i \sigma_{xy} q[Q_0] - g_m L_\omega \oint dx\, {\rm tr}(\partial_x Q_0)^2, \tag{2.54}$$

where the σ_{ij} represent the ''conductances''

$$\sigma_{xx} = f_{xx}(\omega\xi^2) \approx O(\omega\xi^2), \quad \sigma_{xy} = f_{xy}(\omega\xi^2) \approx m + O(\omega^2\xi^4). \tag{2.55}$$

Here, $g_m = m/2$ is the quantized 1D conductance of the chiral edge states, and $L_\omega = m/16\pi\omega\rho_{\rm edge}$ is the frequency-induced length scale. In the limit $\omega \to 0$ the Q_0 entering Eq. (2.54) is forced to obey not only the classical equations of motion on the edge [Eq. (2.53)], but also those arising from the bulk kinetic term in [Eq. (2.54)]. The solutions are known as *instantons*, and just as has happened before in the trivial example with a density of states gap in the bulk, $S_{\rm eff}$ is immaterial as long as $\sigma_{xx} = 0$ and $\sigma_{xy} =$ integer. Therefore, the quantum Hall effect can be understood in terms of a continuous symmetry which is dynamically restored in the limit of large length scales.

In summary we can say that the ''quantum Hall effect'' is a robust and general feature of the instanton vacuum theory for all values of N_r. Our theory of topological quantum numbers is based on two general assumptions only, namely, the existence of a mass gap in the bulk as well as massless excitations at the edge. Both are valid for the σ model in two dimensions for all (non-negative) values of N_r.

The results of this section can be used to demonstrate that a phase transition must occur when σ^0_{xy} passes through half-integer values (or the instanton parameter θ passes through π). The argument[11] is based on the fact that the Hall conductance σ_{xy} must make an integer step when σ^0_{xy} is approached from the integer sides. These phase transitions separate the different instanton vacua which are now labeled by macroscopic quantum numbers (i.e., $\sigma_{xy} =$ integer) and they are distinct from each other by the number of massless modes that exist near the edge of the system. Apart from the close contact with quantum Hall physics, the argument for a phase separation between the different instanton vacua proceeds along similar lines as 't Hooft's duality argument.[23]

Finally, we mention that the results of this paper have interesting consequences for the idea of having a first order phase transition at $\theta = \pi$ (as found, e.g., in the large N theory of the CP^N model[24]). First-order instabilities provide an alternative physical scenario of Hall quantization, and will be discussed elsewhere.[25]

III. DERIVATION OF THE FULL EDGE THEORY

A. Preliminaries

From now on we turn to the fermionic path integral. Following I, one can formulate a complete theory of Q-matrix fields that includes external potentials as well as interactions by making use of such concepts as ''smallness,'' $\mathcal{F}$ invariance, and $\mathcal{F}$ algebra. We will proceed by summarizing the main ingredients of the fermionic path integral approach (Secs. III A and III B). In Sec. III C we present the main steps of a derivation of Q-field theory at the edge, assuming that the Fermi energy lies in a Landau gap. The various manipulations closely follow the effective action procedure for free electrons, and we refer to the original works of Refs. 8 and 11 for the missing details.

1. Notation

Let us start by writing down the Q-field theory for disordered electrons in $2+1$ dimensions in the presence of Coulomb interactions and external potentials, derived in I:[26]

$$S[A,\tilde{Q},\lambda] = -\frac{1}{2g} {\rm Tr}\, \tilde{Q}^2 + {\rm Tr}\ln[i\omega + i\hat{A}_0 + i\hat{\lambda} + \mu - \hat{\mathcal{H}} + i\tilde{Q}] - \frac{1}{2}\beta \int d^2x\, d^2x'\, \lambda^\dagger(x) U_0^{-1}(x-x')\lambda(x'). \tag{3.1}$$

The symbols appearing in this action have the following meaning: $\tilde{Q}(\vec{x})$ is an infinite-dimensional matrix field with two replica indices and two Matsubara frequency indices. (In the derivation of the above action, it arises as a quadratic expression in the original electron field ψ; The saddle point is given by $\tilde{Q}^{\alpha\beta}_{nm} \propto \psi^{\alpha}_{n}\overline{\psi}^{\beta}_{m}$.) Upper Greek indices denote a replica channel, running from 1 to N_r, while lower Latin indices stand for Matsubara frequencies, running from $-\infty$ to ∞. The matrix field $\tilde{Q}$ can be split into ''transverse'' and ''longitudinal'' components:

$$\tilde{Q} = T^{-1} P T, \quad P = P^{\dagger}, \quad T \in \mathrm{SU}(2\tilde{N}). \tag{3.2}$$

Here P has only block-diagonal components in frequency space (i.e., $P^{\alpha\beta}_{nm} \neq 0$ only for $\omega_m \omega_n > 0$), and T is a unitary rotation. The size of the $\tilde{Q}$ matrix is given by $2\tilde{N}$, namely the number of replicas times the size of Matsubara frequency space. The matrix ω is unity in replica space, while in frequency space it is a diagonal containing the fermionic frequencies,

$$\omega^{\alpha\beta}_{nm} = \delta^{\alpha\beta}\delta_{nm}\omega_n, \quad \omega_n = \frac{2\pi}{\beta}\left(n + \frac{1}{2}\right), \tag{3.3}$$

with β the inverse temperature. Tr denotes a matrix trace as well as spatial integration. All spatial integrals are taken in the upper half-plane $y > 0$. The sample edge is given by the line $y = 0$. The $U_0^{-1}(\vec{x} - \vec{x}')$ is the matrix inverse of the Coulomb interaction $U_0(\vec{x} - \vec{x}')$. A_μ is the external potential, and λ is the plasmon field. It is assumed that these fields do not have a static ($n = 0$) component. The ''hat'' notation appearing in Eq. (3.1) is defined as

$$\hat{x} = \sum_{\alpha=1}^{N_r} \sum_{n=-\infty}^{\infty} x^{\alpha}_{n} \tilde{I}^{\alpha}_{n}, \tag{3.4}$$

where $\tilde{I}^{\alpha}_{n}$ is the unity matrix in the αth replica channel, shifted by n places in frequency space:

$$(\tilde{I}^{\alpha}_{n})^{\beta\gamma}_{kl} = \delta^{\alpha\beta}\delta^{\alpha\gamma}\delta_{k-l,n}. \tag{3.5}$$

$\mathcal{H}$ is the kinetic-energy (differential) operator,

$$\mathcal{H} = \frac{1}{2m_e}(\vec{\pi} - \vec{A}) \cdot (\vec{\pi} - \vec{A}), \quad \vec{\pi} = \frac{1}{i}\nabla - \vec{A}_{\mathrm{st}},$$

$$\bar{\pi} = -\frac{1}{i}\bar{\nabla} - \vec{A}_{\mathrm{st}} \tag{3.6}$$

where $\vec{A}_{\mathrm{st}}$ describes the static magnetic field according to $\nabla \times \vec{A}_{\mathrm{st}} = B_{\mathrm{st}}$.

2. Flux-charge composites

In order to describe the fractional quantum Hall effect, one also needs to include a *statistical* or Chern-Simons gauge field a_μ in Eq. (3.1) as follows:

$$S[A, \tilde{Q}, \lambda] \to S[A + a, \tilde{Q}, \lambda] + \frac{i}{8p\pi}\int d\tau\, d^2x\, \varepsilon^{\mu\nu\sigma} a_\mu \partial_\nu a_\sigma, \tag{3.7}$$

with $\varepsilon^{\mu\nu\kappa}$ the antisymmetric tensor in $2+1$ dimensions, and $2p$ an even integer denoting the number of elementary flux quanta h/e attached to every electron. Note that in this procedure the zero-frequency components of all fields are to be treated at a mean-field level. This amounts to adding an extra contribution $\vec{a}_{\mathrm{st}}$ to the static part of the external vector potential $\vec{A}_{\mathrm{st}}$, resulting in an effective magnetic field $B_{\mathrm{eff}} = \nabla \times (\vec{A}_{\mathrm{st}} + \vec{a}_{\mathrm{st}}) = B_{\mathrm{st}} + 2pn_e h/e$, with n_e the mean electron density. Jain's composite fermion mapping is then implemented by integrating out the field a_μ. In this paper, however, we only consider the integer quantum Hall effect; we deal with the fractional effect in a subsequent publication.

3. Landau gap

A theory for the edge is obtained by choosing the chemical potential μ approximately halfway between Landau energies, where the bulk density of states is virtually zero if the disorder is not too strong. The saddle-point equation for $\tilde{Q}$ is given by

$$\tilde{Q}_{\mathrm{sp}} \propto \rho T^{-1} \Lambda T \tag{3.8}$$

where ρ is the density of states and Λ is the matrix appearing in Eq. (2.1), but now with full frequency dependence:

$$\Lambda^{\alpha\beta}_{kl} = \delta^{\alpha\beta}\begin{bmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{bmatrix}_{kl}. \tag{3.9}$$

Since we are interested in the limit $\rho \to 0$, we may replace the full expression for $\tilde{Q}$ [Eq. (3.2)] by a much simpler one:

$$\tilde{Q} \to \varepsilon T^{-1} \Lambda T, \quad \varepsilon \ll 1. \tag{3.10}$$

From detailed earlier work[11] we know that Eqs. (3.2) and (3.10) give rise to identical results as long as the bulk density of states ρ can be safely taken to zero. However, in order to deal with the complications of U(1) gauge invariance (Sec. III B), there is considerable advantage in working with the simplified expression (3.10), and we will refer to the details of more elaborate analyses only when necessary.

B. Gauge invariance and truncation of frequency space

The electromagnetic U(1) gauge transformations in this theory are generated by the $\tilde{I}$ matrices. Multiplication of these matrices is very simple,

$$\tilde{I}^{\alpha}_{n}\tilde{I}^{\beta}_{m} = \delta^{\alpha\beta}\tilde{I}^{\alpha}_{n+m}, \tag{3.11}$$

and they form an abelian algebra. Gauge transformations are given by

$$A_\mu \to A_\mu + \partial_\mu \chi, \quad \tilde{Q} \to e^{i\hat{\chi}} \tilde{Q} e^{-i\hat{\chi}} \tag{3.12}$$

with $\chi^{\alpha}_{0} = 0$. The gauge invariance of Eq. (3.1) is easily checked by writing the transformed Tr ln in the form $\mathrm{Tr}\ln(e^{-i\hat{\chi}}[\cdots]e^{i\hat{\chi}})$, noting that

$$e^{-i\hat{\chi}}\omega e^{i\hat{\chi}} = \omega - \widehat{\partial_0 \chi}, \quad \widehat{\partial_0 \chi} = -i\sum_{n\alpha} \nu_n \chi^{\alpha}_{n} \tilde{I}^{\alpha}_{n} \tag{3.13}$$

and

$$e^{-i\hat{\chi}}(\vec{\pi} - \hat{\vec{A}} - \nabla\hat{\chi})e^{i\hat{\chi}} = \vec{\pi} - \hat{\vec{A}}. \tag{3.14}$$

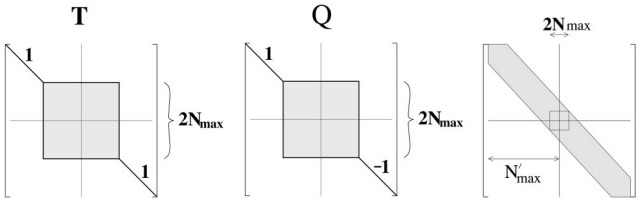

FIG. 5. The truncated matrices T and Q; Also drawn is the frequency band in which $\mathrm{tr}\, I_n^\alpha Q \neq 0$.

In order to facilitate the expansion of the Tr ln term in Eq. (3.1), we make use of the freedom in choosing χ (3.12) as follows. Introducing the notation

$$\tilde{W}=\exp\left(\sum_\alpha \sum_{n\neq 0} \frac{(A_0+\lambda)_n^\alpha}{\nu_n}\tilde{I}_n^\alpha\right), \quad \vec{z}_n^\alpha=\vec{A}_n^\alpha - i\frac{\nabla(A_0+\lambda)_n^\alpha}{\nu_n}, \tag{3.15}$$

with $\nu_n=2\pi n/\beta$, and the quantity $\tilde{R}=\tilde{W}\tilde{Q}\tilde{W}^{-1}$, then the Tr ln can be written as

$$\mathrm{Tr}\ln\left[i\omega+\mu-\frac{1}{2m_e}(\hat{\vec{\pi}}-\hat{\vec{z}})\cdot(\vec{\pi}-\hat{\vec{z}})+i\tilde{R}\right]. \tag{3.16}$$

As was the case in I, we have to impose a cutoff on the size of Matsubara frequency space. Instead of being infinite, all matrices are now of size $2N'_{\max}\times 2N'_{\max}$ in frequency space. The Matsubara indices are restricted to lie in the interval$(-N'_{\max},\ldots,N'_{\max}-1)$.

The truncated version of the $\tilde{I}$ matrices is denoted by I_n^α. The hat notation is now defined with respect to the truncated matrices I_n^α. These no longer span an abelian algebra; instead their commutators are given by

$$(I_n^\alpha I_m^\beta)_{kl}^{\mu\nu}=(\tilde{I}_n^\alpha\tilde{I}_m^\beta)_{kl}^{\mu\nu}g_{l+m},$$

$$[I_n^\alpha,I_m^\beta]_{kl}^{\mu\nu}=\delta^{\alpha\beta\mu\nu}\delta_{k-l,m+n}(g_{l+m}-g_{l+n}), \tag{3.17}$$

where $\delta^{\alpha\beta\mu\nu}$ means that all replica indices have to be the same, and g_i is a step function equal to one if $i\in\{-N'_{\max},\ldots,N'_{\max}-1\}$, and zero otherwise.

In order to handle the U(1) gauge invariance of the theory a second cutoff $N_{\max}\ll N'_{\max}$ is introduced for the matrix field T. With the truncated T we define the truncated equivalent of $\tilde{Q}$ (see Fig. 5),

$$Q=T^{-1}\Lambda T. \tag{3.18}$$

It was shown in Paper I that most of the problems caused by the change from Eq. (3.11) to Eq. (3.17) are avoided by our introduction of the second cutoff. A remnant of the U(1) symmetry is kept in this way: invariance of the action under the truncated equivalent of Eq. (3.12):

$$A_\mu\to A_\mu+\partial_\mu\chi, \quad Q\to e^{i\hat{\chi}}Qe^{-i\hat{\chi}}. \tag{3.19}$$

We do not have a full symmetry of the theory, however, since the integration measure of the truncated field variables Q with finite $N_{\max}$ is *not* invariant under Eq. (3.19). Only in the limit of $N_{\max}\to\infty$ is full symmetry obtained. It is always understood implicitly that this limit is taken in the end.

By introducing the second cutoff $N_{\max}$ we also restrict the interval of the frequency indices on λ_n^α and $(A_\mu)_n^\alpha$ to $n\in[-2N_{\max}+1,2N_{\max}-1]$ (see Fig. 5). This interval corresponds to $\mathrm{tr}\, I_n^\alpha Q\neq 0$.

We now specify the truncated version of Eqs. (3.15) and (3.16) as follows. Write

$$A'_\mu=A_\mu+\delta_{\mu 0}\lambda, \quad W=\exp\left[\sum_{n\alpha} I_n^\alpha(A'_0)_n^\alpha/\nu_n\right],$$

$$R=WQW^{-1}, \tag{3.20}$$

then the action (3.1), (3.10) can be written as (up to a constant)[26]

$$S[Q,\lambda,A]=-\frac{1}{2}\beta\int d^2x\, d^2x'\lambda^\dagger(\vec{x})U_0^{-1}(\vec{x}-\vec{x}')\lambda(\vec{x}')$$
$$+\mathrm{Tr}\ln\left[i\omega+\mu-\frac{1}{2m_e}(\hat{\vec{\pi}}-\hat{\vec{z}})\cdot(\vec{\pi}-\hat{\vec{z}})+i\varepsilon R\right], \tag{3.21}$$

with $\vec{z}$ defined according to

$$\vec{z}_n^\alpha=\vec{A}_n^\alpha-i\nabla(A'_0)_n^\alpha/\nu_n, \tag{3.22}$$

and with ν_n the bosonic frequency $2\pi n/\beta$.

C. Expansion of the Tr ln

1. The quantity X

Let us look at the last term in Eq. (3.21), $X=\mathrm{Tr}\ln[i\omega+\mu-\mathcal{H}_z+i\varepsilon R]$. Introducing the notation

$$D_\omega=TW^{-1}\omega WT^{-1}, \quad D_{\vec{z}}=TW^{-1}\left(\frac{1}{i}\nabla-\hat{\vec{z}}\right)WT^{-1}, \tag{3.23}$$

(where $D_{\vec{z}}$ is not a differential operator), we can write

$$X=\mathrm{Tr}\ln\left[iD_\omega+\mu+i\varepsilon\Lambda\right.$$
$$\left.-\frac{1}{2m_e}(\hat{\vec{\pi}}\cdot\vec{\pi}+\hat{\vec{\pi}}\cdot D_{\vec{z}}+D_{\vec{z}}\cdot\vec{\pi}+D_{\vec{z}}^2)\right]. \tag{3.24}$$

Expansion to first order in D_ω and $D_{\vec{z}}$ yields

$$X\approx\mathrm{Tr}\ln G_0^{-1}+i\,\mathrm{Tr}\,G_0D_\omega-\frac{1}{2m_e}\mathrm{Tr}[G_0\hat{\vec{\pi}}\cdot D_{\vec{z}}+G_0D_{\vec{z}}\cdot\vec{\pi}], \tag{3.25}$$

where G_0 is the bare Green's function $[\mu-(1/2m_e)\hat{\vec{\pi}}\cdot\vec{\pi}+i\varepsilon\Lambda]^{-1}$. The Green's function can be expressed in terms of the eigenfunctions φ_{nj} of the bare Hamiltonian $\mathcal{H}_0=(1/2m_e)\hat{\vec{\pi}}\cdot\vec{\pi}$,

$$\langle x|G_0|x\rangle=\sum_{nj}\frac{|\varphi_{nj}(x)|^2}{\mu-E_{nj}+i\varepsilon\Lambda}, \tag{3.26}$$

$$\left\langle x\left|\frac{G_0\tilde{\pi}+\tilde{\pi}G_0}{2m_e}\right|x\right\rangle = \sum_{nj}\frac{\varphi_{nj}^*\frac{1}{i}\nabla\varphi_{nj}-\varphi_{nj}\frac{1}{i}\nabla\varphi_{nj}^*-2\varphi_{nj}^*\varphi_{nj}\vec{A}}{2m_e(\mu-E_{nj}+i\varepsilon\Lambda)}. \tag{3.27}$$

Using the general relation $\rho(x)=-(1/\pi)\mathrm{Im}\,G^+(x,x)$ for the density of states at the Fermi energy μ, we obtain

$$\langle x|G_0|x\rangle = -i\pi\rho(x)\Lambda + c(x)\mathbf{1},$$

$$\left\langle x\left|\frac{G_0\tilde{\pi}+\tilde{\pi}G_0}{2m_e}\right|x\right\rangle = -i\pi\vec{J}(x)\Lambda+\vec{c}(x)\mathbf{1}, \tag{3.28}$$

where $\vec{J}(x)$ is the current density per energy at the Fermi energy. c and $\vec{c}$ are real functions that disappear from the last two traces in Eq. (3.25). We can now write X in the form

$$\begin{aligned} X &\approx \mathrm{Tr}\ln G_0^{-1}+\pi\int d^2x\,\rho(x)\mathrm{tr}\,\Lambda D_\omega \\ &\quad + i\pi\int d^2x\,\vec{J}(x)\cdot\mathrm{tr}\,\Lambda D_{\vec{z}} \\ &= \mathrm{Tr}\ln G_0^{-1}+\pi\int d^2x\,\rho(x)\mathrm{tr}\,\omega R \\ &\quad + i\pi\int d^2x\,\vec{J}(x)\cdot\mathrm{tr}\left[\frac{1}{i}\Lambda W^{-1}T\nabla(T^{-1}W)-\hat{\vec{z}}R\right]. \end{aligned} \tag{3.29}$$

Since μ lies in a gap, the density of states and the current density are nonzero only at the edge. This means that the surface integral becomes a line integral. Since ρ and $\vec{J}$ are constant on the edge, the resulting expression for X is

$$X \approx \mathrm{Tr}\ln G_0^{-1}+\pi\rho_{\mathrm{edge}}\oint dx\,\mathrm{tr}\,\omega R - i\frac{m}{2}\oint dx\,\mathrm{tr}\,\hat{z}_x R + mS_{\mathrm{top}}[R], \tag{3.30}$$

where we have used that $j_{\mathrm{edge}}=m/2\pi$ with the plateau-center filling fraction $m=(n_e/B)(h/e)$ integer valued, and S_{top} is the topological action

$$S_{\mathrm{top}}[R]=\frac{1}{8}\mathrm{Tr}\,\varepsilon^{ij}R\partial_i R\partial_j R. \tag{3.31}$$

2. The quantity X_0

Equation (3.30), however, is not yet the complete answer. This can be seen from a different expansion procedure which can be followed in the special case where $T=\mathbf{1}$ and $W=\mathbf{1}$. In this case we have, instead of Eq. (3.24),

$$X_0=\mathrm{Tr}\ \ln[i\omega+\mu-\mathcal{H}_z+i\varepsilon\Lambda]\approx\mathrm{Tr}\ \ln G^{-1} - \frac{1}{2}\mathrm{Tr}\left[\frac{G(\tilde{\pi}\cdot\hat{\vec{z}}+\hat{\vec{z}}\cdot\tilde{\pi})}{2m_e}\right]^2-\frac{1}{2m_e}\mathrm{Tr}\,\hat{z}^2G, \tag{3.32}$$

$$G_{nm}=G_n\delta_{nm}=\delta_{nm}\left[i\omega_n+\mu-\frac{1}{2m_e}\tilde{\pi}\cdot\tilde{\pi}\right]^{-1}.$$

This expression can be written as

$$\begin{aligned} X_0 &\approx \mathrm{Tr}\ \ln G^{-1}-\frac{1}{2}\sum_{ij}\sum_{n\alpha}\int d^2x\,d^2x'(z_i)_n^\alpha(x)(z_j)_{-n}^\alpha(x') \\ &\quad\times(\Pi_{ij})_n^\alpha(x,x') \\ &\quad -\frac{1}{2m_e}\sum_i\sum_{n\alpha}\int d^2x(z_i)_n^\alpha(z_i)_{-n}^\alpha\,\mathrm{tr}\,G(x,x). \end{aligned} \tag{3.33}$$

The ''polarization operator'' Π_{ij} is given by

$$\begin{aligned} (\Pi_{ij})_n^\alpha(x,x') &= \left(\frac{1}{2m_e}\right)^2\mathrm{tr}[G(x,x')(\tilde{\pi}_i+\tilde{\pi}_i) \\ &\quad\times I_n^\alpha G(x',x)(\tilde{\pi}_j+\tilde{\pi}_j)I_{-n}^\alpha] \\ &= \left(\frac{1}{2m_e}\right)^2\sum_k G_{k+n}(x,x')(\tilde{\pi}_i+\tilde{\pi}_i) \\ &\quad\times G_k(x',x)(\tilde{\pi}_j+\tilde{\pi}_j). \end{aligned} \tag{3.34}$$

The frequency sum can be split in two parts: (i) k and $k+n$ have the same sign; (ii) k and $k+n$ have opposite signs. Case (ii) has been done in great detail in the context of the self-consistant Born approximation. The conclusion is that (ii) does not contribute either to σ_{xx} or σ_{xy} when μ is in a density of states gap. Case (i) for $i\neq j$, using the relation $\tilde{\pi}+\tilde{\pi}=-i2m_e[G^{-1},\vec{x}]$, gives rise to the familiar ''Streda'' form for σ_{xy}. For $i=j$, the last two contributions in Eq. (3.33) sum up to zero. We arrive at the following expression:

$$X_0\approx\mathrm{Tr}\ \ln G^{-1}+\frac{1}{2}m\sum_{n\alpha}\int d^2x\,n\vec{z}_n^\alpha\times\vec{z}_{-n}^\alpha. \tag{3.35}$$

3. Matching X and X_0

Now we have to find a match between the first-order result (3.30) for $T\neq\mathbf{1}$, $W\neq\mathbf{1}$ and the second-order result (3.35) for $T=\mathbf{1}$ and $W=\mathbf{1}$. This match can be written in two equivalent ways.[26]

$$\begin{aligned} &\mathrm{Tr}\ \ln G_0^{-1}+\pi\rho_{\mathrm{edge}}\oint dx\,\mathrm{tr}\,\omega R \\ &\quad + m\left(\frac{1}{8}\varepsilon^{ij}\,\mathrm{Tr}\,R[D_i,R][D_j,R]-\frac{i}{2}\mathrm{Tr}\,R\nabla\times\hat{\vec{z}}\right) \\ &= \mathrm{Tr}\ \ln G_0^{-1}+\frac{m}{2v_d}\oint dx\,\mathrm{tr}\,R(\omega-iv_d\hat{z}_x) \\ &\quad + mS_{\mathrm{top}}[R]-\frac{im\beta}{4\pi}\int d^2x\,\vec{z}^\dagger\times\partial_0\vec{z}, \end{aligned} \tag{3.36}$$

with v_d the electron drift velocity at the edge:

$$v_d=m/(2\pi\rho_{\mathrm{edge}}). \tag{3.37}$$

Notice that the l.h.s. of Eq. (3.36) represents the lowest order terms in a formal series expansion of the Tr ln [Eq. (3.21)] in powers of ω and the covariant derivative $D_j=\partial_j$

$-i\tilde{A}_j$. On the other hand, the r.h.s. clearly matches the results denoted by X [Eq. (3.30)] and X_0 [Eq. (3.35)]. It is important to remark that the equality in Eq. (3.36) is a direct consequence of the peculiar structure of $\mathcal{F}$ algebra [Eq. (3.37)].

We next employ $\mathcal{F}$ algebra in order to express the result Eq. (3.36) in terms of the matrix field variable Q rather than R. The following relations can be derived[26]

$$\begin{aligned}
\mathrm{tr}\,\omega R &= \mathrm{tr}\,\omega Q + \mathrm{tr}\,\hat{A}_0' Q - \frac{\beta}{2\pi} A_0'^{\dagger} A_0', \\
\mathrm{tr}\,R\hat{z}_x &= \mathrm{tr}\,Q\hat{A}_x - \mathrm{tr}\,Q\partial_0^{-1}(\partial_x \hat{A}_0') - \frac{\beta}{\pi} A_x^{\dagger} A_0' \\
&\quad + \frac{\beta}{\pi} A_0'^{\dagger}\partial_0^{-1}(\partial_x A_0'), \\
S_{\mathrm{top}}[R] &= S_{\mathrm{top}}[Q] - \frac{i}{2}\oint dx\,\mathrm{tr}\,Q\partial_0^{-1}(\partial_x \hat{A}_0') \\
&\quad + \frac{i\beta}{4\pi}\oint dx A_0'^{\dagger}\partial_0^{-1}(\partial_x A_0'), \\
\beta\int d^2x\,\vec{z}^{\dagger}\times\partial_0\vec{z} &= -\beta\int d^2x\,\varepsilon^{\mu\nu\kappa}(A_\mu')^{\dagger}(\partial_\nu A_\kappa') \\
&\quad -\beta\oint dx[\partial_0 A_x - \partial_x A_0']^{\dagger}\partial_0^{-1}A_0'.
\end{aligned} \tag{3.38}$$

Eqs. (3.36)–(3.38) lead to the following final result for the action [Eq. (3.21)].

$$\begin{aligned}
S[Q,A,\lambda] &= S_c[\lambda] + S_b[\lambda,A] + S_Q[Q,\lambda,A], \\
S_c &= -\frac{1}{2}\beta\int d^2x\,d^2x'\lambda^{\dagger}(x)U_0^{-1}(\vec{x}-\vec{x}')\lambda(x'), \\
S_b &= \frac{im\beta}{4\pi}\int d^2x\,\varepsilon^{\mu\nu\kappa}(A_\mu')^{\dagger}(\partial_\nu A_\kappa') - \frac{m\beta}{4\pi v_d}\oint dx\,A_0'^{\dagger}A_-', \\
S_Q &= \frac{m}{2v_d}\oint dx\,\mathrm{tr}\,Q(\omega+\hat{A}_-') + mS_{\mathrm{top}}[Q].
\end{aligned} \tag{3.39}$$

The first term is the Coulomb energy contribution from the plasmon field; S_b is a ''boson'' action (this adjective will become clear later on); the last expression S_Q contains the action for the Q field and the coupling of Q with λ and A. We have defined a ''minus'' direction as

$$A_-' = A_0' - iv_d A_x, \tag{3.40}$$

reflecting the chirality inherent in the problem.

IV. CHIRAL EDGE BOSONS

In this section we take the theory one step further, and derive the theory of chiral edge bosons, similar to the one obtained by Wen[1] in a phenomenological approach to abelian quantum Hall states. For noninteracting electrons such a formulation is readily obtained (Sec. IV A). For interacting electrons, however, the procedure is more complicated, and we first derive an effective Finkelstein-type action of the Q field at the edge, obtained by eliminating the plasmon field λ (Sec. IV B 1). In Sec. IV B 2 we show that the theory provides complete information on the response of the system to external fields. We derive an edge anomaly for the interacting electron gas, and show the connection with Laughlin's gauge argument. The complete theory for interactions as well as the $2+1$-dimensional Chern-Simons theory are derived in Sec. IV B 3. In Sec. IV C we give some explicit results on the single-particle Green's function which enters the expression for electron tunneling into the quantum Hall edge. This, then, completes the theory of the integral quantum Hall edge.

A. Noninteracting case

In the case of free electrons, only the fields Q and A are present in Eq. (3.39). In order to obtain an effective action for A_μ we integrate out Q. We make use of Eq. (2.22) with $2\omega\to\omega_n$, and write the two point function as follows:

$$\langle \mathrm{tr}\,I_n^{\alpha} Q(-q)\,\mathrm{tr}\,I_{-n}^{\alpha} Q(q)\rangle = \frac{\beta v_d}{2\pi m}\frac{\omega_n}{\omega_n + iv_d q}. \tag{4.1}$$

After elementary algebra we obtain the result

$$S = \frac{im\beta}{4\pi}\left[\int d^2x\,\varepsilon^{\mu\nu\kappa}A_\mu^{\dagger}\partial_\nu A_\kappa + \oint dx\,E_x^{\dagger}\partial_-^{-1}A_-\right], \tag{4.2}$$

with the following meaning of the symbols:

$$E_x = \partial_0 A_x - \partial_x A_0, \quad \partial_- = \partial_0 - iv_d\partial_x, \tag{4.3}$$

and the inverse ∂_-^{-1} is given by

$$\begin{aligned}
(\partial_-^{-1}F)(x,\tau) &= \frac{1}{(2\pi)^2}\int dx'\,d\tau' F(x',\tau')\int_{-\infty}^{\infty}dk\int_{-\infty}^{\infty}d\omega \\
&\quad\times\frac{\exp[ik(x-x')-i\omega(\tau-\tau')]}{-i\omega+v_d k} \\
&= \frac{1}{2\pi}\int dx'\,d\tau' F(x',\tau') \\
&\quad\times\left[\frac{\theta(\tau-\tau')}{v_d(\tau-\tau')-i(x-x')+\eta}\right. \\
&\quad\left.+\frac{\theta(\tau'-\tau)}{v_d(\tau-\tau')-i(x-x')-\eta}\right],
\end{aligned} \tag{4.4}$$

with F an arbitrary function, θ the step function and η a regulator. The operation ∂_-^{-1} does not commute with ∂_-. It is easily checked that $\partial_-(\partial_-^{-1}F)=F$, but on the other hand we have

$$\partial_-^{-1}(\partial_- F) = F - F^{\mathrm{res}}, \tag{4.5}$$

with F^{res} defined as that part of F which satisfies $\partial_- F = 0$. Another property of this operation is

$$\int dx\,d\tau F_1(\partial_-^{-1}F_2) = -\int dx\,d\tau(\partial_-^{-1}F_1)F_2. \tag{4.6}$$

Eqs. (4.5) and (4.6) can be used for the purpose of demonstrating the $U(1)$ gauge invariance of the action (4.2).

Notice that the same result, Eq. (4.2), can also be obtained more directly from Eq. (3.39) if one makes use of the statement of gauge invariance in truncated frequency space, Eq. (3.19). More specifically, the χ field can be chosen in such a way that the Q and A_μ fields in Eq. (3.39) decouple

$$\partial_- \chi = -A_-, \quad \chi = -\partial^{-I} A_- + \chi^{\rm res}$$

The χ dependence now enters the theory through the expression $-\oint dx E_x^\dagger \chi = \oint dx E_x^\dagger \partial^{-1} A_-$, which precisely equals the second term in Eq. (4.2).

Action (4.2) can also be written as a path integral over m charge 1 bosons,

$$S[A,\varphi_i] = \frac{i\beta}{4\pi}\sum_{i=1}^{m}\left[\int d^2x\, \varepsilon^{\mu\nu\kappa} A_\mu^\dagger \partial_\nu A_\kappa - \oint dx (D_x \varphi_i^\dagger D_- \varphi_i - E_x^\dagger \varphi_i)\right], \tag{4.8}$$

where the covariant derivative D is defined as $D_\mu \varphi_i = \partial_\mu \varphi_i - A_\mu$.

As a general remark, we can say that our introduction of m chiral boson fields is obviously not a unique procedure as long as one limits oneself to the charge sector of the theory alone. In oder to show that the theories of Eq. (3.39) and Eq. (4.8) are completely equivalent representations of the quantum Hall edge dynamics, it is obviously necessary to extend the analysis to include such quantities like the tunneling denisty of states (Sec. IV C), heat transport and the specific heat. The later will be reported elsewhere.[25] In order to make contact with Ref. 1 we mention that Eq. (4.8) is equivalent to a Chern-Simons bulk theory with m gauge fields g^i that represent potentials for the electron currents, coupled to the external potentials A_μ,

$$S[A,g^i] = \frac{i\beta}{4\pi}\sum_{i=1}^{m}\int d^2x\, \varepsilon^{\mu\nu\kappa}[-(g_\mu^i)^\dagger \partial_\nu g_\kappa^i + 2(g_\mu^i)^\dagger \partial_\nu A_\kappa], \tag{4.9}$$

where the g^i have the gauge fixing constraint $g_-^i|_{\rm edge} = 0$. In Appendix B we explicitly show how integration over the potentials g^i leads to Eq. (4.8).

B. Coulomb case

1. Integration over λ and Q

Now we look at the full action (3.39). In this expression the plasmon field λ is contained in the following way:

$$S_b[A'] = S_b[A] + \frac{im\beta}{2\pi}\int d^2x\, \lambda^\dagger B - \frac{m\beta}{4\pi v_d}\oint dx(\lambda^\dagger\lambda + 2\lambda^\dagger A_0),$$
$$S_Q[A'] = S_Q[A] + \frac{m}{2v_d}\oint dx\, {\rm tr}\,\hat{\lambda} Q. \tag{4.10}$$

Integrating out the plasmon field λ, we obtain an effective action for Q coupled to A, which we organize as follows:

$$S = S_0[Q] + S_{\rm int}[Q,A] + S_b[A] + S_{\rm flux}[A]. \tag{4.11}$$

The first term is given by

$$S_0[Q] = mS_{\rm top}[Q] + S_{\rm F}[Q] - \frac{m\pi}{4\beta}\sum_{n\alpha}\int \frac{dk_x}{2\pi}\frac{1}{v^{\rm eff}(k_x)}|{\rm tr}\, {\rm I}_n^\alpha Q|^2, \tag{4.12}$$

with

$$S_{\rm F}[Q] = \frac{m\pi}{4\beta v_d}\oint dx\left[\sum_{n\alpha}{\rm tr}\, I_n^\alpha Q\, {\rm tr}\, I_{-n}^\alpha Q + 4\,{\rm tr}\,\eta Q\right], \tag{4.13}$$

the edge analog of the $\mathcal{F}$-invariant Finkelstein action for the bulk [I], and

$$v^{\rm eff}(k_x) = v_d + mU_0(k_x), \tag{4.14}$$

the ''effective velocity,'' where $U_0(k_x) = (2\pi)^{-1}\int dk_y U_0(\vec{k})$ is the Coulomb interaction on the edge. The last term in Eq. (4.12) is the edge version of the ''Coulomb'' term from I. Note that the Finkelstein and Coulomb terms together can be written as

$$\frac{m}{2v_d}\oint dx\,{\rm tr}\,\omega Q + \frac{m\pi}{4\beta v_d}\sum_{n\alpha}\int\frac{dk_x}{2\pi}\frac{\rho_{\rm edge}}{U_0^{-1}(k_x)+\rho_{\rm edge}}|{\rm tr}\, I_n^\alpha Q|^2, \tag{4.15}$$

where the expression in front of the $|{\rm tr}\, IQ|^2$ is just the 1D screened Coulomb interaction. The other terms in Eq. (4.11) are a coupling term

$$S_{\rm int}[Q,A] = \frac{m}{2}\int\frac{dk_x}{2\pi}\frac{1}{v^{\rm eff}(k_x)}{\rm tr}\, Q\hat{A}_c^{\rm eff}, \tag{4.16}$$

a ''boson'' term

$$S_b[A] = \frac{im\beta}{4\pi}\int d^2x\varepsilon^{\mu\nu\kappa}(A_\mu^{\rm eff})^\dagger\partial_\nu A_\kappa^{\rm eff} - \frac{m\beta}{4\pi}\int\frac{dk_x}{2\pi}\frac{1}{v^{\rm eff}(k_x)}(A_0^{\rm eff})^\dagger A_c^{\rm eff}, \tag{4.17}$$

and a flux-flux interaction term

$$S_{\rm flux}[A] = \frac{\beta}{2}\left(\frac{m}{2\pi}\right)^2\int d^2x\, d^2x' B^\dagger(\vec{x})U_0(\vec{x}-\vec{x}')B(\vec{x}'). \tag{4.18}$$

Here we have introduced an ''effective'' gauge field which contains a Coulomb correction to the scalar potential,

$$\vec{A}^{\rm eff} = \vec{A}, \quad A_0^{\rm eff}(\vec{x}) = A_0(\vec{x}) + \frac{im}{2\pi}\int d^2x' U_0(\vec{x}-\vec{x}')B(\vec{x}'), \tag{4.19}$$

and an effective ''minus'' direction denoted by the subscript c

$$\partial_c = \partial_0 - iv^{\rm eff}\partial_x; \quad A_c = A_0 - iv^{\rm eff}A_x. \tag{4.20}$$

Comparing result (4.11) with the free particle case [Eq. (3.39) without λ], we see that the presence of the Coulomb interaction has the following effects: (i) the appearance of the flux-flux interaction term $S_{\rm flux}[A]$ and of the screened Coulomb interaction in $S_0[Q]$, (ii) the replacements $A_0 \to A_0^{\rm eff}$ and $A_- \to A_c^{\rm eff}$, and (iii) the replacement $v_d \to v^{\rm eff}(k_x)$. For what follows, it is convenient to rewrite the first three terms of Eq. (4.11) as

$$\begin{aligned} S_0+S_{\rm int}+S_b=&\frac{im\beta}{4\pi} \\ &\times\left[\int d^2x\,\varepsilon^{\mu\nu\kappa}(A_\mu^{\rm eff})^\dagger\partial_\nu A_\kappa^{\rm eff}-\oint dx\,A_x^\dagger A_c^{\rm eff}\right] \\ &+mS_{\rm top}[Q]+S_{\rm F}[Q] \\ &-\frac{m\pi}{4\beta}\sum_{n\alpha}\int\frac{dk_x}{2\pi}\frac{1}{v^{\rm eff}(k_x)} \\ &\times\left|{\rm tr}\,I^\alpha_{-n}Q-\frac{\beta}{\pi}(A_c^{\rm eff})^\alpha_n\right|^2, \end{aligned} \tag{4.21}$$

where, as in the bulk (see paper I), the gauge field couples to Q only via the gauge non-invariant ''Coulomb'' term in Eq. (4.12). However, compared to the bulk case where the coupling results in the gauge-invariant combination $[{\rm tr}\,IQ-(\beta/\pi)A_0]$, the situation is more subtle in the edge case. The expression $[{\rm tr}\,IQ-(\beta/\pi)A_c^{\rm eff}]$ appearing in Eq. (4.21) is, in fact, gauge variant, but this gauge variance is exactly what one needs to compensate for the edge contributions resulting from gauge transformations of the ''boson'' action S_b and the topological term. Therefore, the complete action [Eq. (4.1)$+S_{\rm flux}$] is fully gauge invariant.

We now proceed as in Sec. IV A and integrate out the Q field. This is done in the same way as for the noninteracting case: either by doing it directly or by choosing a guage χ such that Q decouples from A_μ, i.e.,

$$\partial_c\chi=-A_c^{\rm eff}. \tag{4.22}$$

The only difference lies in the fact that we now work with effective quantities. We then obtain the effective action for the external field A_μ in the presence of Coulomb interactions:

$$\begin{aligned} S[A]=&\frac{im\beta}{4\pi}\left[\int d^2x\,\varepsilon^{\mu\nu\kappa}(A_\mu^{\rm eff})^\dagger\partial_\nu A_\kappa^{\rm eff}\right. \\ &\left.+\oint dx(\partial_c^{-1}A_c^{\rm eff})^\dagger E_x^{\rm eff}\right] \\ &+\frac{\beta}{2}\left(\frac{m}{2\pi}\right)^2\int d^2x\,d^2x'B^\dagger(\vec{x})U_0(\vec{x}-\vec{x}')B(\vec{x}'). \end{aligned} \tag{4.23}$$

Again, the difference from the free-particle case is the appearance of a flux-flux term and various replacements by effective quantities.

2. Edge currents and Laughlin's gauge argument

Action (4.23) contains complete information on the response of the system to external electromagnetic fields. We define the current as $j^\mu(\vec{x})=\delta S/\delta A_\mu(\vec{x})$. In this way we find

$$j^0(\vec{x})=\frac{im}{2\pi}[B-\delta(y)\partial_c^{-1}E_x^{\rm eff}], \tag{4.24}$$

$$\begin{aligned} j^1(\vec{x})=&-\frac{im}{2\pi}\left[E_y-\partial_y\int d^2x'\,U_0(\vec{x}-\vec{x}')j^0(\vec{x}')\right] \\ &-\frac{mv_d}{2\pi}\delta(y)\partial_c^{-1}E_x^{\rm eff}, \end{aligned} \tag{4.25}$$

$$j^2(\vec{x})=\frac{im}{2\pi}\left[E_x-\partial_x\int d^2x'\,U_0(\vec{x}-\vec{x}')j^0(\vec{x}')\right]. \tag{4.26}$$

It is easily verified that $\partial_\mu j^\mu=0$. The edge currents are obtained by taking only those terms that possess a δ function. On the edge we obtain

$$j^0_{\rm edge}=-\frac{im}{2\pi}\partial_c^{-1}E_x^{\rm eff}, \tag{4.27}$$

$$j^1_{\rm edge}=-iv_d\cdot j^0_{\rm edge}. \tag{4.28}$$

For the edge anomaly, this yields

$$\partial_\mu j^\mu_{\rm edge}(x)=-\frac{im}{2\pi}\left[E_x-\partial_x\int d^2x'\,U_0(x,\vec{x}')j^0(\vec{x}')\right]. \tag{4.29}$$

By applying Laughlin's gauge argument,[10] one can now directly relate the conductances defined by the bulk and the edge. For example, let us do a linear response calculation for the case where N flux quanta h/e are created somewhere inside a hole in the sample. The charge q flowing from one edge into the other is found using Eq. (4.29),

$$dq/d\tau=-i\oint\partial_\mu j^\mu_{\rm edge}=\frac{m}{2\pi}d\Phi/d\tau, \tag{4.30}$$

where Φ is the total flux Nh/e enclosed by the contour integral. This yields $q=mN$, as it should.

3. Interacting chiral bosons

As was the case in the free-electron situation, we can write theory (4.23) as an edge boson coupled to the external field, exactly of the form of Eq. (4.8), but now with effective quantities and an extra flux-flux term:

$$\begin{aligned} S[A,\varphi_i]=&\frac{i\beta}{4\pi}\sum_{i=1}^m\left[\int d^2x\,\varepsilon^{\mu\nu\kappa}(A_\mu^{\rm eff})^\dagger\partial_\nu A_\kappa^{\rm eff}\right. \\ &\left.-\oint dx(D_x\varphi_i^\dagger D_c^{\rm eff}\varphi_i-\varphi_i^\dagger E_x^{\rm eff})\right] \\ &+\frac{\beta}{2}\left(\frac{m}{2\pi}\right)^2\int d^2x\,d^2x'B^\dagger(\vec{x})U_0(\vec{x}-\vec{x}')B(\vec{x}'). \end{aligned} \tag{4.31}$$

As in the noninteracting case, this result is equivalent to a Chern-Simons bulk theory of the form of Eq. (4.9). In this case the action for the electron currents is given by

$$S[A,g_i]=\frac{i\beta}{4\pi}\sum_{i=1}^{m}\int d^2x\,\varepsilon^{\mu\nu\kappa}[-(g_\mu^i)^\dagger\partial_\nu g_\kappa^i+2(g_\mu^i)^\dagger\partial_\nu A_\kappa^{\rm eff}]+\frac{\beta}{2}\left(\frac{m}{2\pi}\right)^2\int d^2x\,d^2x'B^\dagger(\vec{x})U_0(\vec{x}-\vec{x}')B(\vec{x}'),\tag{4.32}$$

with the gauge fixing conditions

$$\left[g_-^j(k_x)-i\frac{m}{\sqrt{2\pi}}U_0(k_x)\sum_{a=1}^{m}g_x^a(k_x)\right]_{\rm edge}=0.\tag{4.33}$$

It is very instructive also to write Eq. (4.31) in the following way:

$$S=-\frac{\beta}{2}\left(\frac{1}{2\pi}\right)^2\sum_{i,j=1}^{m}\int d^2x\,d^2x'U_0(\vec{x}-\vec{x}')\nabla\times[\theta(y)\vec{D}\varphi_i(\vec{x})]^\dagger\nabla'\times[\theta(y')\vec{D}\varphi_j(\vec{x}')]+\frac{i\beta}{4\pi}\sum_{i=1}^{m}\left[\int d^2x\,\varepsilon^{\mu\nu\kappa}A_\mu^\dagger\partial_\nu A_\kappa-\oint dx(D_x\varphi_i^\dagger D_-\varphi_i-E_x^\dagger\varphi_i)\right].\tag{4.34}$$

Notice that there are no effective quantities in this expression: the Coulomb interaction is completely contained in the first term. The charge density is given by $(m/2\pi)[B+\delta(y)m^{-1}\Sigma_i D_x\varphi_i]$. Notice also that we have written a two-dimensional integral containing φ_i, even though the boson fields only exist on the edge. This is not a problem, since the φ_i only are evaluated at the edge.

C. Tunneling density of states

In I we expressed the one-particle Green's function which enters the tunneling density of states in terms of the matrix Q variable as follows:

$$\langle Q^{\alpha\alpha}(\tau_2,\tau_1,\vec{x}_0)\rangle=\sum_{n=-\infty}^{\infty}e^{i\nu_n(\tau_2-\tau_1)}\langle Q_{nn}^{\alpha\alpha}(\vec{x}_0)\rangle.\tag{4.35}$$

The gauge transformation, that in Eq. (4.21) decouples Q from A, introduces an extra factor into the path integral over Eq. (4.23):

$$\exp\{-i[[\partial_c^{-1}A_c^{\rm eff}]^\alpha(\tau_2,\vec{x}_0)-[\partial_c^{-1}A_c^{\rm eff}]^\alpha(\tau_1,\vec{x}_0)]\}.\tag{4.36}$$

When decoupling the quadratic edge term in A [Eq. (4.23)] with the use of boson fields, this factor translates to

$$\exp\left\{-i\int_{\tau_1}^{\tau_2}d\tau\partial_\tau\varphi_j^\alpha(\tau,\vec{x}_0)\right\},\quad j=1,\ldots,m\tag{4.37}$$

in Eq. (4.31). The decoupling is not a unique procedure, since combinations of the boson fields φ_i can be chosen other than Eq. (4.37). However, the above form is the only one that yields the fermionic exponent for the expectation value $\langle Q\rangle$:

$$\left\langle\exp\left\{-i\int_{\tau_1}^{\tau_2}d\tau\,\partial_\tau\varphi_i^\alpha(\tau,\vec{x}_0)\right\}\right\rangle\propto(\tau_2-\tau_1)^{-S},\quad S=1.\tag{4.38}$$

(See Appendix A for the explicit calculation.) Notice that we would have had a serious problem at this point if we had not excluded the zero-momentum components of φ_i^α when we introduced these auxiliary fields. A redefinition of the integration measure, $\int\mathcal{D}\varphi\to\int\mathcal{D}[\varphi+f]$, with $\partial_x f(x,\tau)=0$, would yield a result depending on the arbitrary function f.

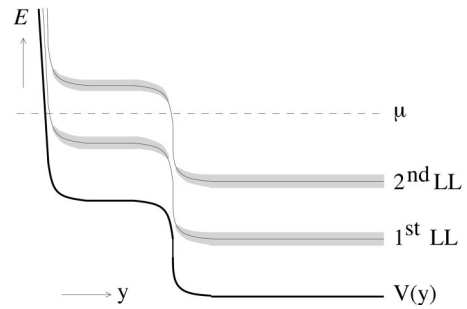

FIG. 6. Spatially separated edge channels.

V. LONG-RANGE DISORDER

In Sec. II B we introduced the idea of percolating edge states as a model for smooth, slowly varying randomness. Application of Q-field theory then provides an effective and elegant way of describing the transport properties of the network model near the percolation threshold. In this section we extend the network theory of percolating edge states in several ways. We show that the Coulomb interactions can dramatically alter the behavior of the electron gas, depending on the physical process that one is interested in. The concept of a ''tunneling density of states,'' that describes the tunneling of electrons into the quantum Hall edge, is particularly sensitive to the presence of long-range electron-electron interactions. In Sections V A–V C we derive an ''effective'' theory of chiral edge bosons that includes the effect of Coulomb interactions between the edge and bulk electrons. This leads to a tunneling exponent S that varies continuously with the filling fraction ν like $1/\nu$. This result is in dramatic contrast to the Fermi liquid predictions of Sec. IV C which apply to isolated edges alone. We start out (Sec. V A) with the chiral boson formulation of the network model, and employ the Laughlin gauge argument in order to illustrate the fundamental differences between transport and edge tunneling (Sec. V B). Section V C describes one of the most important aspects of this section. It deals with the detailed mechanism by which the ''neutral'' modes are eliminated from the effective theory for edge tunneling. We end this section with a computation of the inelastic relaxation rate (Sec. V D) that enters into the transport problem at finite temperatures (Sec. II B 4).

A. Separation of edge channels

Long-range disorder can cause the edge states of different Landau levels to become spatially separated. A potential

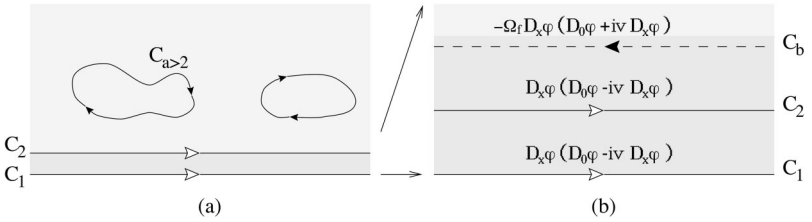

FIG. 7. (a) Plot of equipotential contours corresponding to filling fraction $\nu=2-\varepsilon$. (b) Effective edge theory for the filling fraction $\nu=2-\varepsilon$. The dashed line represents the (anti)chiral contribution from the bulk orbitals.

fluctuation at the edge can lift all states in such a way that new ''edge'' states are created (see Fig. 6). If the chemical potential lies between the shifted and unshifted energy of a Landau level, the edge states of this Landau level will be situated inside the sample, not on the outermost edge. If there are several potential jumps of this kind, all the edge channels can become separated. They can also start wandering into the interior of the sample.

We propose that ''edge channel separation'' is the dominant effect of smooth potential fluctuations as opposed to ''interchannel scattering,'' which only occurs when the potential changes abruptly. In this section we wish to embark on the problem of smooth potential fluctuations in the presence of the Coulomb interactions.

In order to fix the thought we imagine a quantum Hall sample with filling fraction $\nu=2-\varepsilon$. Figure 7(a) illustrates the equipotential contours. We may distinguish between the localized (closed) orbitals in the bulk of the sample and the extended (chiral) edge states.

This picture leads us to the idea of describing the chiral bosons by *one* field $\varphi(\vec{x})$ that lives on all the ''edges'' instead of independent fields for every edge. Action (4.31) then becomes

$$S=\frac{i\beta}{4\pi}\Bigg[\int d^2x\, n(\vec{x})\varepsilon^{\mu\nu\kappa}A^{\dagger}_{\mu}\partial_{\nu}A_{\kappa} - \sum_{a=1}^{M} s_a \oint_{C_a} dx(D_x\varphi^{\dagger}[D_0\varphi - is_a v_d D_x\varphi]-E_x^{\dagger}\varphi)\Bigg] - \frac{\beta/2}{(2\pi)^2}\int d^2x\, d^2x' U_0(\vec{x}-\vec{x}')\nabla\times[n(\vec{x})\vec{D}\varphi(\vec{x})]^{\dagger} \nabla'\times[n(\vec{x}')\vec{D}\varphi(\vec{x}')]. \tag{5.1}$$

The n is a function of position labeling the ''local'' filling fraction: outside the sample $n(\vec{x})$ is zero; going inward, it increases by one every time you cross an edge, until it reaches its bulk value m. At the bulk orbitals, $n(\vec{x})$ jumps again. [In the case $\nu=2-\varepsilon$, depicted in Fig. 7(a), $n(\vec{x})=1$ inside the closed orbitals.]

Each edge is described by a contour labeled C_a, with $a=1,\ldots,m$ for the edge states and $a=m+1,\ldots,M$ for the closed bulk orbitals. The coordinate x, appearing in the edge terms, is defined on the contour and is taken in the positive (anticlockwise) direction. The symbol s_a,

$$s=(\overbrace{+1,\cdots,+1}^{m},-1,\cdots,-1) \tag{5.2}$$

incorporates the fact that the contours with $a\leqslant m$ and $a>m$ carry opposite current and charge densities. For simplicity we take the drift velocity v_d the same for all edges. Integrating out the boson field yields the generalization of Eq. (4.23):

$$S[A]=\frac{i\beta}{4\pi}\Bigg[\int d^2x\, n(\vec{x})\varepsilon^{\mu\nu\kappa}(A^{\rm eff}_{\mu})^{\dagger}\partial_{\nu}A^{\rm eff}_{\kappa} + \sum_a s_a \oint_{C_a} dx[\partial_c^{-1}(\partial_x A^{\rm eff}_c)-A_x]^{\dagger}A^{\rm eff}_c\Bigg] + \frac{\beta}{2}\left(\frac{1}{2\pi}\right)^2\int d^2x\, d^2x' n(\vec{x})B^{\dagger}(\vec{x}) \times U_0(\vec{x}-\vec{x}')n(\vec{x}')B(\vec{x}'). \tag{5.3}$$

The notation ∂_c (at contour C_a) now has the sign s_a in front of the velocity and contains Coulomb interactions with *all* contours instead of just C_a itself. The definition of the ''effective'' potential $A_0^{\rm eff}$ has also slightly changed:

$$A_0^{\rm eff}(\vec{x})=A_0(\vec{x})+\frac{i}{2\pi}\int d^2x' U_0(\vec{x}-\vec{x}')n(\vec{x}')B(\vec{x}'). \tag{5.4}$$

For completeness, in Appendix D we also present the generalization of the action $S[Q,A]$ [Eq. (4.11)] for the case of separated edge channels. Note that we are addressing the situation where the chemical potential is away from the narrow ''percolation'' regime indicated by W_0 in Fig. 4. We will next exploit the simplicity of our model and demonstrate that the Hall conductance and the tunneling density of edge states are fundamentally different quantities that correspond to completely different physical processes.

B. Hall conductance

First it is straightforward to generalize the results of Sec. IV B 2 to include the separated edge channels and the bulk states into Laughlin's flux argument. Differentiating action (5.3) with respect to A_{μ}, we obtain the generalized form of the currents [(4.24)–(4.26)],

$$j^0(\vec{x})=\frac{i}{2\pi}\Bigg[n(\vec{x})B-\sum_{a=1}^{M} s_a\delta(\vec{x}\ \text{ on }\ C_a)\partial_c^{-1}E_x^{\rm eff}\Bigg], \tag{5.5}$$

$$j^i(\vec{x})=-i\frac{n(\vec{x})}{2\pi}\varepsilon^{ij}\Bigg[E_j-\partial_j\int d^2x' U_0(\vec{x}-\vec{x}')j^0(\vec{x}')\Bigg] - \frac{v_d}{2\pi}\sum_{a=1}^{M}\delta(\vec{x}\ \text{ on }\ C_a)\partial_c^{-1}E_x^{\rm eff}(\vec{e}_{\parallel}^{\,a})_i, \tag{5.6}$$

where the vector $\vec{e}_{\parallel}^{a}$ is tangent to the contour C_a and points in the positive direction. Again it is easy to check that $\partial_\mu j^\mu = 0$, i.e., that current conservation is respected. The edge currents are given by

$$j^0_{\rm edge}(C_a) = -\frac{i}{2\pi} s_a \partial_c^{-1} E_x^{\rm eff}, \tag{5.7}$$

$$j^x_{\rm edge}(C_a) = -i s_a v_d j^0_{\rm edge}(C_a). \tag{5.8}$$

The edge anomaly applies to each bulk orbital and edge state separately:

$$\partial_\mu j^\mu_{\rm edge}(C_a) = -\frac{i}{2\pi} s_a \left[E_x - \partial_x \int d^2x' U_0(x,\vec{x}') j^0(\vec{x}') \right]. \tag{5.9}$$

As expected, the sign s_a determines whether charge is transported into an edge or from an edge into the bulk. By repeating Laughlin's flux argument it is now demonstrated explicitly that the localized bulk orbitals do not affect the transport of charge from one sample edge to the other, independent of the electron-electron interactions; taking Eq. (5.9) and performing the contour integral over C_a, we obtain the charge transported per unit of time from the ath channel,

$$dQ_a/d\tau = -i \oint_{C_a} dx\, \partial_\mu j^\mu_{\rm edge} = s_a \frac{1}{2\pi} d\Phi_a/d\tau, \tag{5.10}$$

where Φ_a is the magnetic flux enclosed by C_a. For $a > m$ this flux is obviously zero, since the localized bulk orbitals do not encircle the hole in the sample. This, then, shows that the Hall conductance is quantized (equal to m) independent of ε.

C. Tunneling density of states

Laughlin's flux argument for the Hall conductance expresses the quantum Hall state as an exact ''excited'' state of the system. Tunneling processes into the edge, on the other hand, are expressed in terms of eigenstates near the Fermi energy, i.e., the tunneling density of states, and due to the Coulomb interactions this quantity is sensitive to the presence of bulk orbitals. We start from action (5.1), omitting the replica indices for notational simplicity and putting $A_\mu = 0$:

$$\begin{aligned} S = &-\frac{i}{4\pi} \int d\tau \sum_{j=1}^{M} s_j \oint_{C_j} dx\, \partial_x \varphi (\partial_0 \varphi - i s_j v_d \partial_x \varphi) \\ &- \frac{1}{8\pi^2} \int d\tau \sum_{j,j'=1}^{M} s_j s_{j'} \oint_{C_j} dx \\ &\times \oint_{C_{j'}} dx' \partial_x \varphi U_0(x,x') \partial_{x'} \varphi. \end{aligned} \tag{5.11}$$

Following Sec. IV C, Eq. (4.37), the one-particle Green's function can be written as

$$G(\tau_2 - \tau_1) = \langle \exp -i[\varphi(\tau_2, x_0) - \varphi(\tau_1, x_0)] \rangle, \tag{5.12}$$

where x_0 denotes a point on the edge contour C_1. The presence of the Coulomb interactions makes the computation of G a complicated two-dimensional problem. Some procedure needs to be found which extracts the lowest energy excitations from Eq. (5.11), We follow the strategy of taking the boson fields as a two-dimensional field variable, and we then collect the terms with smallest momenta. This procedure is done in position space and we proceed by giving the details of a step by step analysis. The results for the tunneling exponents are given at the end of Sec. V C 4, which also contains a brief summary in the end.

1. Gradient expansion

The interaction term in Eq. (5.11), can be written as a sum over area integrals,

$$\begin{aligned} &-\frac{1}{8\pi^2} \int d\tau \sum_{j,j'=1}^{M} \int_{C_j} d^2x \int_{C_{j'}} d^2x' s_j s_{j'} \partial_a \varphi(\vec{x}) \\ &\quad \times U_{ab}(\vec{x} - \vec{x}') \partial_b \varphi(\vec{x}'), \end{aligned} \tag{5.13}$$

with

$$U_{ab}(\vec{x} - \vec{x}') = \varepsilon_{ac} \varepsilon_{bd} \frac{\partial}{\partial x_c} \frac{\partial}{\partial x'_d} U_0(\vec{x} - \vec{x}'). \tag{5.14}$$

Since we are only interested in the φ with the smallest momenta, we can make the replacement

$$\sum_{j=m+1}^{M} \int_{C_j} d^2x \to \Omega_f \int_{C_b} d^2x. \tag{5.15}$$

Ω_f stands for the fraction of the total area that is enclosed by all the bulk orbitals together. The contour C_b is not sharply defined, and is located somewhere close to the edge [see Fig. 7(b)]. It encloses the region within which the bulk orbitals are contained. The joint Coulomb effects of the bulk orbitals will effectively be comprised on this contour. For the terms in Eq. (5.11) containing $\partial_x \varphi \partial_0 \varphi$, we can write

$$\begin{aligned} \sum_{j>m} \oint_{C_j} dx\, \partial_x \varphi \partial_0 \varphi &= \sum_{j>m} \int_{C_j} d^2x \nabla \times (\nabla \varphi \partial_0 \varphi) \\ &\to \Omega_f \int_{C_b} d^2x\, \nabla \times (\nabla \varphi \partial_0 \varphi) \\ &= \Omega_f \oint_{C_b} dx\, \partial_x \varphi \partial_0 \varphi. \end{aligned} \tag{5.16}$$

The expression $\Sigma_{j>m} \oint_{C_j} dx (\partial_x \varphi)^2$ averages out to $\kappa \int_{C_b} d^2x (\nabla \varphi)^2$ with κ some positive constant related to the total length of all the bulk contours. If there are substantial stretches where a bulk orbital runs along the edge, interaction terms will arise, leading to a term $\oint_{C_b} dx (\partial_x \varphi)^2$.

Note that in doing replacement (5.15) in Eq. (5.13), one also needs to introduce correction terms that compensate for the errors made when the separation $|\vec{x} - \vec{x}'|$ is ''small'' (of the order of the average size of the orbitals or less) and U_{ab} does not vary slowly. These corrections are of the form $\int d^2x (\nabla \varphi)^2$.

Then there are also extra correction terms that will arise if there are regions where a bulk orbital runs along the edge.

This correction takes the form of a short-ranged interaction between C_b and all the other contours (including C_b).

Having done replacement (5.15) and writing the interaction terms again as contour integrals, we have the action

$$S=-\frac{i}{4\pi}\int d\tau\left[\sum_{j=1}^{m}\oint_{C_j}dx\,\partial_x\varphi\partial_-\varphi-\varepsilon\oint_{C_b}dx\,\partial_x\varphi\partial_0\varphi\right]$$
$$-\frac{1}{8\pi^2}\int d\tau\left[\sum_{jj'=1}^{m}\oint_{C_j}dx\oint_{C_{j'}}dx'\,\partial_x\varphi U\partial_{x'}\varphi\right.$$
$$+\varepsilon^2\oint_{C_b}dx\,dx'\,\partial_x\varphi(U+V_b)\partial_{x'}\varphi$$
$$\left.-2\varepsilon\sum_{j=1}^{m}\oint_{C_j}dx\oint_{C_b}dx'\,\partial_x\varphi(U+V_j)\partial_{x'}\varphi\right]$$
$$-g\int d\tau\int_{C_b}d^2x(\nabla\varphi)^2, \tag{5.17}$$

where g is a positive constant. We have identified Ω_f with ε, since the fraction of the area occupied by bulk states is exactly the deviation from integer filling. We have written $V_b(x,x')$ for the short-ranged interaction between two points on C_b; $V_j(x,x')$ denotes the short-ranged interaction between a point x on C_j and a point x' on C_b. The precise expression for V is unknown due to the fact that it has its origin in the twilight zone near the edge, where it is unclear whether a term contributes to the bulk or edge action. Comparing this result (5.17) with Eq. (5.11), we see that the presence of the interacting bulk states effectively leads to the appearance of an additional (anti)chiral boson on the contour C_b, an extra short-ranged interaction with this contour, and a lower-dimensional leftover bulk term $\int(\nabla\varphi)^2$.

2. Effect of the bulk term

In order to be able to calculate the tunneling density of states [Eq. (5.12)], we need an effective theory for the edge degrees of freedom, and therefore we have to understand how they are affected by the left over bulk term. To this end, we are going to split bulk and edge degrees of freedom. We write the bulk term as $\int_{C_b}d^2x(\nabla\Phi)^2$, where Φ represents the bulk degrees of freedom and is treated as an integration variable independent of φ. To reflect the fact that it is actually an extension of φ into the bulk, we impose some boundary condition on Φ, for instance $\Phi|_{\text{edge}}=\varphi$ or $\partial_\perp\Phi|_{\text{edge}}=\partial_\perp\varphi$. ($\partial_\perp$ is the derivative perpendicular to the contour.) The effect of the bulk term on the edge theory is obtained by integrating out Φ, which leads to an effective action for the boundary conditions. Let us consider a general scenario and impose the boundary conditions $\Phi|_{\text{edge}}=\psi_0$ and $\partial_\perp\Phi|_{\text{edge}}=\psi_1$, using constraint multipliers k_0 and k_1, respectively:

$$e^{S_{\text{eff}}[\psi_0(x),\psi_1(x)]}=\int\mathcal{D}[\Phi(\vec{x})]\mathcal{D}[k_0(x)]\mathcal{D}[k_1(x)]$$
$$\times\exp\left\{i\oint dx\,k_0(\Phi-\psi_0)\right.$$
$$\left.+i\oint dx\,k_1(\partial_\perp\Phi-\psi_1)-g\int d^2x(\nabla\Phi)^2\right\}. \tag{5.18}$$

For notational simplicity we have omitted the time dependence and the subscript C_b under all the integrals. We first wish to integrate Eq. (5.18) over $\Phi(\vec{x})$, keeping k_0 and k_1 fixed. For this purpose we split Φ, which has free boundary values, into a bulk and an edge part by writing

$$\Phi=\Phi_L+\hat{\Phi},\quad \partial_\perp\Phi_L|_{\text{edge}}=\partial_\perp\Phi|_{\text{edge}}\quad \partial_\perp\hat{\Phi}|_{\text{edge}}=0 \tag{5.19}$$

where Φ_L satisfies Laplace's equation

$$\nabla^2\Phi_L(\vec{x})=0. \tag{5.20}$$

$\Phi_L(\vec{x})$ is completely determined by $\partial_\perp\Phi_L$ on the edge, which we now take as an independent edge degree of freedom denoted by $E_1(x)$. Introducing the 2D Green's function G,

$$G(\vec{x},\vec{x}')=\frac{1}{2\pi}\ln|\vec{x}-\vec{x}'|,\quad \nabla^2G(\vec{x},\vec{x}')=\delta(\vec{x}-\vec{x}'), \tag{5.21}$$

and using Green's theorem, we solve Laplace's equation and obtain for $\Phi_L(\vec{x})$,

$$\Phi_L(\vec{x})=-\oint dx'\left[G(\vec{x},x')E_1(x')\right.$$
$$\left.-\Phi_L(x')\frac{\partial G}{\partial y'}(x,y;x',0)\right]. \tag{5.22}$$

This expression tells us that we need to know Φ_L on the edge in order to evaluate Φ_L in the bulk. Luckily, we do not need the full 2D $\vec{x}$ dependence, since due to the splitting [Eq. (5.19)], Φ_L will be evaluated at the edge only. Using a special property of the Green's function [Eq. (5.21)], namely, $[\partial_{y'}G](x,0;x',0)=0$, we can explicitly write Φ_L on the edge as a function of E_1:

$$\Phi_L(x)=-\oint dx'G(x,x')E_1(x'). \tag{5.23}$$

The action, written in terms of $\hat{\Phi}$ and E_1, is now given by

$$S=-g\int d^2x(\nabla\hat{\Phi})^2-g\oint\oint E_1GE_1+2g\oint dx\,E_1\hat{\Phi}$$
$$-i\oint\oint Gk_0+i\oint dx\,k_0(\hat{\Phi}-\psi_0)$$
$$+i\oint dx\,k_1(E_1-\psi_1), \tag{5.24}$$

where we have used the shorthand notation $\oint\oint AGB$ for the expression $\oint dx\oint dx'A(x)G(x,x')B(x')$. Integrating out $\hat{\Phi}$ is now simply done by replacing $\hat{\Phi}$ by its saddle-point value. Varying the action with respect to $\hat{\Phi}$, keeping E_1 fixed, we obtain the saddle-point equation

$$\nabla^2\hat{\Phi}+\delta(y)\left[E_1+\frac{i}{2g}k_0\right]=0. \tag{5.25}$$

Using the Green's function's property $[\partial_{y'}G](x,0;x',0)=0$ again, we find the following solution on the edge:

$$\hat{\Phi}(x) = -\oint dx' G(x,x') \left[E_1 + \frac{i}{2g} k_0 \right](x'). \quad (5.26)$$

In substituting this solution into Eq. (5.24), we do not need the full 2D $\vec{x}$ dependence of $\hat{\Phi}(\vec{x})$, since we can write $\int d^2x (\nabla\hat{\Phi})^2 = -\int d^2x\, \hat{\Phi}\nabla^2\hat{\Phi}$ and $\nabla^2\hat{\Phi}$ is an expression restricted to the edge. Substitution of Eq. (5.26) into Eq. (5.24) yields

$$S = -2g \oint\oint E_1 G E_1 + \frac{1}{4g} \oint\oint k_0 G k_0 - i \oint dx\, k_0 \psi_0 - 2i \oint\oint G E_1 + i \oint dx\, k_1 (E_1 - \psi_1). \quad (5.27)$$

Integrating out k_0 is straightforward, and gives

$$S = g \oint\oint (\psi_0 G^{-1} \psi_0 + 2 E_1 G E_1) + 4g \oint dx\, \psi_0 E_1 + i \oint dx\, k_1 (E_1 - \psi_1). \quad (5.28)$$

In the end we integrate out k_1, yielding the constraint $E_1 = \psi_1$. The final result for $S_{\rm eff}[\psi_0, \psi_1]$ becomes

$$\begin{aligned} S_{\rm eff}[\psi_0, \psi_1] &= g \oint\oint (\psi_0 G^{-1} \psi_0 + 2 \psi_1 G \psi_1) + 4g \oint dx\, \psi_0 \psi_1 \\ &= g \oint\oint (\psi_0, \psi_1) \begin{pmatrix} G^{-1} & 2 \\ 2 & 2G \end{pmatrix} \begin{pmatrix} \psi_0 \\ \psi_1 \end{pmatrix}. \end{aligned} \quad (5.29)$$

We are going to put $\psi_0 = 0$ in order to avoid double counting of $(\partial_x \varphi)^2$ terms at the edge, and $\psi_1 = \partial_\perp \varphi$. Action (5.29) becomes

$$S[\partial_\perp \varphi] = 2g \oint\oint \partial_\perp \varphi G \partial_\perp \varphi. \quad (5.30)$$

This edge term, derived from the interaction with the bulk orbitals, is seriously going to affect the tunneling exponent. A quick way to see this is as follows: on the contours $C_1, \ldots, C_b$, the field $\varphi(\vec{x})$ can be written as $\varphi(x,y \text{ on } C_1)$ plus perpendicular derivatives. For the tunneling exponent, only $\varphi|_{C_1}$ is needed, so we can integrate out the perpendicular derivatives in [Eq. (5.17), minus bulk term plus Eq. (5.30)] to obtain an effective action for φ on C_1. The dominant part of the 1D propagator for $\partial_\perp \varphi$ is given by $G^{-1}(k) \propto |k|$, from which it follows that all terms introduced by the integration over $\partial_\perp \varphi$ are irrelevant. Higher powers of $\partial_\perp$ are even less relevant. Replacing all the φ in Eq. (5.17) by $\varphi|_{C_1}$, we obtain a term $\nu \oint dx\, \partial_x \varphi \partial_0 \varphi$, leading to a tunneling exponent $S = 1/\nu$ instead of the free-particle result $S = 1$. In Sec. V C 3 we are going to derive this result more formally, based on a consideration of the neutral modes in the theory where the edge channels are not spatially separated.

3. Demise of the neutral modes; example $\nu = 1 - \varepsilon$

In the long-wavelength limit, the contours $C_1, \ldots, C_b$ are lying so close together that we can effectively return to the picture where all the edge channels are sitting on top of each other. We label the channels $\varphi_1(x), \ldots, \varphi_m(x), \varphi_b(x)$. Let us for simplicity's sake first consider the case $\nu = 1 - \varepsilon$, where we just have the two fields φ_1 and φ_b. In terms of these fields, action (5.17), without the bulk term and the bulk effect (5.30), takes the form (again using abbreviated notation)

$$\begin{aligned} S_0[\varphi_1, \varphi_b] &= -\frac{1}{4\pi} \oint dx [\partial_x \varphi_1 \partial_0 \varphi_1 - \varepsilon \partial_x \varphi_b \partial_0 \varphi_b] \\ &\quad - \frac{1}{8\pi^2} \oint\oint U [\partial_x \varphi_1 - \varepsilon \partial_x \varphi_b]^2 \\ &\quad - \frac{1}{8\pi^2} \sum_{k,l=1,b} \oint dx\, V_{kl} \partial_x \varphi_k \partial_x \varphi_l. \end{aligned} \quad (5.31)$$

We have put all the short-range contributions into the 2×2 velocity matrix V. We next define a ''charged mode'' Γ and a ''neutral mode'' γ in such a way that only the charged mode ''feels'' the long-range part of the interaction;

$$\begin{aligned} \Gamma &= \frac{1}{\nu}(\varphi_1 - \varepsilon \varphi_b), \quad \gamma = \varphi_1 - \varphi_b \\ \varphi_1 &= \Gamma - \frac{\varepsilon}{\nu}\gamma, \quad \varphi_b = \Gamma - \frac{1}{\nu}\gamma. \end{aligned} \quad (5.32)$$

In the basis (Γ, γ) action (5.31) becomes

$$\begin{aligned} S_0[\Gamma, \gamma] &= -\frac{1}{4\pi} \oint dx \left[\nu \partial_x \Gamma \partial_0 \Gamma - \frac{\varepsilon}{\nu} \partial_x \gamma \partial_0 \gamma \right] \\ &\quad - \frac{\nu^2}{8\pi^2} \oint\oint U (\partial_x \Gamma)^2 \\ &\quad - \frac{1}{8\pi^2} \oint dx [\partial_x \Gamma\, \partial_x \gamma] \hat{V} \begin{bmatrix} \partial_x \Gamma \\ \partial_x \gamma \end{bmatrix}, \end{aligned} \quad (5.33)$$

where $\hat{V}$ is the velocity matrix in this basis. The expression $\partial_\perp \varphi$ in the theory for spatially separated channels is in the single-edge picture evidently equivalent to the neutral mode $\gamma \propto \varphi_b - \varphi_1$. The leftover bulk contribution [Eq. (5.30)] therefore translates into an extra term involving the neutral mode;

$$S_{\rm bulk}[\gamma] = \text{const} \oint\oint \gamma G \gamma. \quad (5.34)$$

The tunneling density of states is now expressed as

$$\langle \exp -i\varphi_1 |_{\tau_1}^{\tau_2} \rangle \propto \int \mathcal{D}[\Gamma] \mathcal{D}[\gamma] \exp \left[-i \left(\Gamma - \frac{\varepsilon}{\nu}\gamma \right) \Big|_{\tau_1}^{\tau_2} + S_0[\Gamma, \gamma] + S_{\rm bulk}[\gamma] \right]. \quad (5.35)$$

If we perform the integration over γ first, we see that the bulk part of the action yields the following contribution to the inverse propagator: $G(k) \propto 1/|k|$, which is dominant at

low momenta. The integration over γ yields Γ-Γ terms of order $k^5\hat{V}(k)$. These are clearly irrelevant. For the tunneling density of states we can write

$$\langle \exp -i\varphi_1|_{\tau_1}^{\tau_2}\rangle \propto \int \mathcal{D}[\Gamma]\exp(-i\Gamma|_{\tau_1}^{\tau_2}+S_{\rm eff}[\Gamma])$$

$$S_{\rm eff}[\Gamma]=-\frac{i\nu}{4\pi}\oint dx\,\partial_x\Gamma\partial_0\Gamma-\frac{\nu^2}{8\pi^2}\oint\oint\partial_x\Gamma U\partial_x\Gamma$$

$$-\frac{1}{8\pi^2}\oint dx\,\partial_x\Gamma\hat{V}_{\Gamma\Gamma}\partial_x\Gamma. \tag{5.36}$$

For small momenta the $\hat{V}$ essentially reduces to a constant and we can use the results of appendix A, obtaining

$$\langle \exp -i\varphi_1|_{\tau_1}^{\tau_2}\rangle \propto (\tau_2-\tau_1)^{-S},\quad S=1/\nu. \tag{5.37}$$

4. General case $\nu=m-\varepsilon$

The results for $\nu=1-\varepsilon$ are easily generalized. From the bulk channel φ_b and the edge channels $\varphi_1,\ldots,\varphi_m$ we construct a charged mode γ_0 and m neutral modes $\gamma_1,\ldots,\gamma_m$ as follows:

$$\gamma_0=\frac{1}{\nu}\left(\sum_{k=1}^{m}\varphi_k-\varepsilon\varphi_b\right),$$

$$\gamma_a=\frac{1}{a}\left(\sum_{k=1}^{a}\varphi_k-a\varphi_{a+1}\right),\quad a=1,\ldots,m \tag{5.38}$$

where we define φ_{m+1} as φ_b. The neutral modes $\gamma_1,\ldots,\gamma_{m-1}$ are the usual ones for a theory with m edges. They are mutually perpendicular and normal to the charged mode. The additional γ_m is normal to the other neutral modes but not to the charged mode. The φ's are expressed in terms of the γ's as follows:

$$\varphi_b=\gamma_0-\frac{m}{\nu}\gamma_m,$$

$$\varphi_k=\gamma_0-\frac{\varepsilon}{\nu}\gamma_m-\left(1-\frac{1}{k}\right)\gamma_{k-1}+\sum_{a=k}^{m-1}\frac{1}{a+1}\gamma_a,\quad k\leqslant m. \tag{5.39}$$

Equation (5.31) is generalized to

$$S[\varphi]=-\frac{1}{4\pi}\oint dx\left[\sum_{j=1}^{m}\partial_x\varphi_j\partial_0\varphi_j-\varepsilon\partial_x\varphi_b\partial_0\varphi_b\right]$$

$$-\frac{1}{8\pi^2}\oint\oint U\left[\sum_{j=1}^{m}\partial_x\varphi_j-\varepsilon\partial_x\varphi_b\right]^2$$

$$-\frac{1}{8\pi^2}\sum_{k,l=1}^{m+1}\oint dx V_{kl}\partial_x\varphi_k\partial_x\varphi_l. \tag{5.40}$$

Again, all the short-range contributions have been put into a velocity matrix V, which now has dimension $(m+1)\times(m+1)$. Writing Eq. (5.40) in terms of the γ basis, we obtain

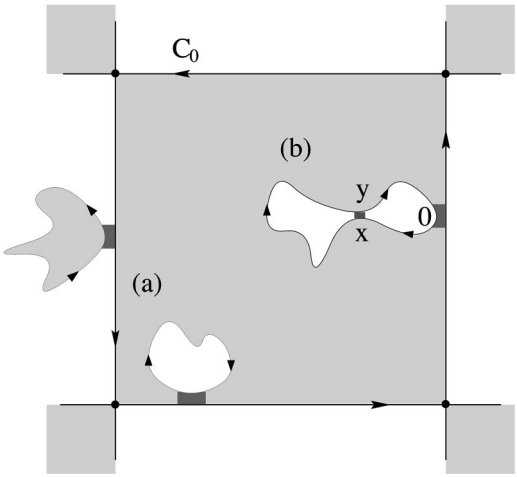

FIG. 8. (a) Pointlike interaction between conducting electrons and the localized electrons at ''nearly saddle points.'' (b) The coordinates 0, x, and y are the points of interaction along the localized contour.

$$S[\gamma]=-\frac{1}{4\pi}\oint dx\left[\nu\partial_x\gamma_0\partial_0\gamma_0+\sum_{a=1}^{m-1}\frac{a}{a+1}\partial_x\gamma_a\partial_o\gamma_a\right.$$

$$\left.-m\frac{\varepsilon}{\nu}\partial_x\gamma_m\partial_0\gamma_m\right]-\frac{\nu^2}{8\pi^2}\oint\oint U(\partial_x\gamma_0)^2$$

$$-\frac{1}{8\pi^2}\sum_{a,c=0}^{m}\oint dx\,\hat{V}_{ac}\partial_x\gamma_a\partial_x\gamma_c, \tag{5.41}$$

where $\hat{V}$ is the velocity matrix in the basis of γ's. The argument of Eqs.(5.34)–(5.37) can be applied again, in a slightly modified form: the neutral modes are equivalent to $\partial_\perp\varphi$ and higher derivatives. (A basis $\hat{\gamma}$ can be found for the neutral modes in which $\hat{\gamma}_n$ corresponds to the 1D lattice discretization of $\partial_\perp^n\varphi$.) On dimensional grounds the propagator for the nth normal derivative of φ has to be proportional to k^{2n-1}, leading to irrelevant contributions. A more concrete way of making this statement would be to generalize the analysis presented in Eqs. (5.18)–(5.29), including boundary conditions for the higher normal derivatives. However, that would also require us to take into account higher order terms in the φ theory [Eq. (5.17)]. The resulting effective action for the charged mode γ_0 is of the form of Eq. (5.36), with $\nu=m-\varepsilon$.

We can summarize the results of Sec. V as follows: We have seen that the Fermi liquid result $S=1$ is obtained for the tunneling density of states (i) when the Coulomb interactions are omitted, or (ii) when interactions are included but only short length scales are considered. An interacting theory for the lowest-lying excitations, which are slowly varying field configurations, yields completely different results. The presence of bulk orbitals, interacting mutually and with the edge states, is effectively described by an extra edge channel with prefactor $-\varepsilon$ plus a remnant of the interactions in the bulk of the form $\int(\nabla\varphi)^2$. The leftover bulk term serves to make all the neutral edge modes irrelevant, yielding an effective edge action for the one remaining, charged, mode. Due to the presence of the extra 'bulk' channel, the prefactor of this effective action $S[\Gamma]$ becomes $m-\varepsilon=\nu$, which is a

continuous parameter in sharp contrast to the integer quantized m. For the tunneling exponent we obtain $S=1/\nu$.

D. Computation of $\tau_{\rm in}$

We next return to the problem of the plateau transitions. Following Sec. II B we expect that the transport at high temperatures is dominated by interactions between the conducting electrons on the backbone saddlepoint network and those on the disconnected pieces or clusters.

The fundamental quantity to compute is the characteristic time $\tau_{\rm in}$ that is needed for the backbone electrons to equilibrate with the rest of the network. In order to set up a theory for relaxation, we consider the ''nearly saddle points'' in the network, where tunneling is not possible but where the Coulomb forces nevertheless produce ''sudden changes'' in the motion of the conducting electrons. Figure 8 illustrates the interaction of the saddle-point network with disconnected orbitals. The ''nearly saddle points'' where the Coulomb forces are most effective are indicated by the shaded areas. We can model the situation by introducing a delta-function potential which acts in the small areas of the ''nearly saddle points'' only. The action can be written as

$$S_{\rm eff}[\varphi]=S[\varphi_0]+\sum_i S[\varphi_i] - \sum_i \int d\tau\, \partial_x\varphi_0(\vec{a}_i)U_i\partial_x\varphi_i(\vec{a}_i), \quad (5.42)$$

where $S[\varphi_0]$ is the action for the chiral boson field on a link of the saddle-point network that we denote as the contour C_0,

$$S[\varphi_0]=\frac{i}{4\pi}\int d\tau \oint_{C_0} dx\, \partial_x\varphi_0\partial_-\varphi_0. \quad (5.43)$$

This contour is taken to be very large or infinite. Similarly, we define chiral boson fields φ_i on the disconnected but large contours C_i;

$$S[\varphi_i]=\frac{i}{4\pi}\int d\tau \oint_{C_i} dx\, \partial_x\varphi_i\partial_+\varphi_i. \quad (5.44)$$

The sum in the interaction term in Eq. (5.42) is over the discrete set of nearly saddle points $\vec{a}_i$ along the contour C_0 where the fields φ_0 and φ_i interact with an appropriate, random strength U_i. This problem is in many ways quite similar to the problem of interacting edge channels with a randomly varying separation between them. We proceed along the same lines as in Ref. 27, and introduce a self-energy Σ for the density-density correlation of the field φ_0. If we denote the Fourier transforms of the propagators $\langle\partial_x\varphi_0(x,\tau)\partial_x\varphi_0(x',\tau')\rangle$ and $\langle\partial_x\varphi_i(x,\tau)\partial_x\varphi_i(x',\tau')\rangle$ (with x and x' parametrizing the positions on the contours C_0 and C_j, respectively) as

$$D_0(\omega,q)=\frac{2\pi}{\beta}\frac{-q}{i\omega-v_dq}, \quad D_j(\omega,q)=\frac{2\pi}{\beta}\frac{-q}{i\omega+v_dq}, \quad (5.45)$$

then the introduction of a self-energy takes the form

$$D_0(\omega,q)\to-\frac{q}{i\omega-(v_d+\Sigma)q}. \quad (5.46)$$

To lowest order in the interaction potential we may write

$$\Sigma(\omega)=-iz\overline{U_j^2}\int\frac{dq}{2\pi}D_j(\omega,q)=\frac{z}{2v_d^2}\overline{U_j^2}|\omega|. \quad (5.47)$$

Here the bar stands for the average over the random positions $\vec{a}_i$ along C_0 and z is the linear density of saddlepoints. Result (5.47) can be used to obtain an expression for $1/\tau_{\rm in}$, i.e., the imaginary part of the self-energy as it appears in the electron Green's function $\mathcal{G}(\omega,q)$ as follows:

$$1/\tau_{\rm in}=\int\frac{d\omega\,dq}{(2\pi)^2}\Sigma(\omega)\mathcal{G}(\varepsilon-\omega,q). \quad (5.48)$$

The $\tau_{\rm in}$ determines the rate at which the electrons on the backbone cluster equilibrate with the rest of the electronic orbitals. We find $\tau_{\rm in}^{-1}\propto\varepsilon^2$ or T^2 at finite temperatures. This admittedly crude approach toward electron relaxation can be improved in several ways. For example, as the most important correction to the self-energy (5.47) we find the self-interacting orbitals as depicted in Fig. 8(b). These corrections replace the momentum integral in Eq. (5.47) in the following way (in space-time notation):

$$\int\frac{dq}{2\pi}D_j(\omega,q) = \int d\tau\, e^{-i\omega(\tau-\tau')}D_j(0,0;\tau-\tau')$$

$$D_j(0,0;\tau-\tau')\to D_j(0,0;\tau-\tau') + \int d\tau_0\int_0^L dx\int_x^L dy \times D_j(0,x;\tau-\tau_0)\tilde{U}_jD_j(y,L;\tau_0-\tau'), \quad (5.49)$$

where x,y are the positions of the ''nearly saddle point'' where the self-interaction takes place. The integrals stand for the averaging over positions and all dimensional factors are absorbed into $\tilde{U}_j$. The length of the orbital is given by L and boundary conditions $x\equiv x+L$ and $y\equiv y+L$ are understood. Equation (5.49) can be rewritten as a shift in the chemical potential:

$$\int dq\,D_j(\omega,q)\to\int dq\,\frac{-iq}{i\omega-\delta\mu+v_dq}, \quad \delta\mu=\tilde{U}_j. \quad (5.50)$$

This leads to a modified self-energy according to

$$\Sigma(\omega)\to\frac{z}{2v_d^2}\overline{U_i^2}(\omega+i\delta\mu)\mathrm{sgn}(\omega). \quad (5.51)$$

The shift $\delta\mu$ can be translated into a shift in the expression for $\tau_{\rm in}^{-1}$ following

$$\tau_{\rm in}^{-1}(\varepsilon,\delta\mu)=\left(1+i\delta\mu\frac{\partial}{\partial\varepsilon}\right)\tau_{\rm in}^{-1}(\varepsilon). \quad (5.52)$$

After the analytic continuation to real energies ($i\varepsilon \to \varepsilon$) has been performed, we obtain the final result $\tau_{\rm in}^{-1} \propto \varepsilon$ or $\tau_{\rm in}^{-1} \propto T$ at finite temperatures. More generally, we expect the equilibration rate to be given by a regular series expansion in powers of T which is dominated by the lowest order $\tau_{\rm in}^{-1} \propto T$ as T approaches absolute zero.

VI. SUMMARY AND CONCLUSIONS

We have shown that massless edge excitations are an integral part of the instanton vacuum theory with free boundary conditions. Massless edge excitations have fundamental consequences for the ''stability'' of topological quantum numbers and for the quantization of the Hall conductances in particular. We have used the formalism of $\mathcal{F}$ algebra, introduced in our previous work, and derived a complete theory of the edge. We have established the fundamental connection between the instanton vacuum and Chern-Simons gauge theory. Both theories have previously been studied independently and with different physical objectives. We have shown that our approach to edge physics enables one to address several long-standing problems of smooth disorder and interaction effects. We have pointed out that fundamental differences exist between tunneling at the edge and electron transport. Transport experiments inject electrons directly into edge states; these electrons do not have enough time to equilibrate with the rest of the sample, and are therefore effectively decoupled from the bulk. A tunneling measurement, however, probes eigenstates of the whole system, which involve not only edge electrons, but also localized bulk orbitals. Since tunneling processes do not probe the incompressibility of the electron gas, they are generally treated incorrectly by the theory of isolated edges. By taking into account the effect of Coulomb interactions between the edge and the localized bulk states, we have derived an effective edge theory that predicts a tunneling exponent $1/\nu$.

For the plateau transitions we have constructed a percolation model of interacting edges. We have shown how inelastic scattering at the ''nearly saddle points'' sets the temperature scale at which the transport coefficients cross over from mean-field behavior to critical scaling. This crossover can involve arbitrarily low temperatures and it explains the lack of scaling in the transport data taken from samples with long-range disorder at finite temperatures. Our mean field expression for the conductances agrees with recent empirical fits to transport data at plateau transitions. The results of this paper serve as the basic starting point for a subsequent one,[14] where we extend the theory to include the statistical gauge fields and the fractional quantum Hall regime.

ACKNOWLEDGMENT

This research was supported in part by INTAS (Grant No. 96-0580).

APPENDIX A: ONE-DIMENSIONAL PROPAGATOR WITH COULOMB INTERACTION

In this appendix we calculate the correlation function $G(\tau,0)$ for the charged boson fields φ_i [Eq. (4.31)]:

$$G(\tau,x) = \langle \varphi_i(\tau,x)\varphi_i(0,0)\rangle, \quad \tau > 0. \tag{A1}$$

In momentum and frequency space this correlator is given by (we omit the label i since it is of no consequence)

$$\langle \varphi_a(k)\varphi_{-b}(-k')\rangle = \frac{2\pi i}{\beta}\frac{\delta_{ab}\delta(k-k')}{k[\omega_a + ikv^{\rm eff}(k)]}. \tag{A2}$$

We write the Coulomb interaction and the effective velocity $v^{\rm eff}$ in the form

$$U_0(k) = -c\sqrt{2\pi}\ln(k/\Lambda)^2, \quad v^{\rm eff}(k) = -mc\ln(k/\Lambda D)^2 \tag{A3}$$

where c is a positive constant indicating the strength of the Coulomb interaction, Λ is an ultraviolet cutoff, and $D = \exp(v_d/2mc)$. We will only consider low momenta $|k| < \lambda\Lambda$, with $\lambda < 1$, so that we are well away from the point where the Hamiltonian becomes negative.

We take the Fourier transform of Eq. (A2) and change the frequency sum to an integral, writing $\Sigma_n \to (\beta/2\pi)\int d\omega$,

$$\begin{aligned}\partial_\tau G(\tau,0) &= \frac{i}{2\pi}\int_{-\lambda\Lambda}^{\lambda\Lambda} dk\, v^{\rm eff}(k)\int_{-\infty}^{\infty}\frac{e^{i\omega\tau}d\omega}{\omega + ikv^{\rm eff}(k)} \\ &= -\int_{-\lambda\Lambda}^{\lambda\Lambda} dk\, v^{\rm eff}(k)\,\theta(-kv^{\rm eff})e^{kv^{\rm eff}(k)\tau}.\end{aligned} \tag{A4}$$

The step function $\theta(-kv^{\rm eff})$ constrains the integration interval to $k<0$. We can split the last expression in Eq. (A4) into two parts, using $\ln k\, dk = d(k\ln k - k)$, and obtain

$$\begin{aligned}\partial_\tau G(\tau,0) = &-\frac{1}{\tau}\left[1 - \left(\frac{\lambda}{D}\right)^{2mc\tau\lambda\Lambda}\right] \\ &- 2mc\Lambda D\int_0^{\lambda/D} du\,\exp[2mc\tau\Lambda D u\ln u].\end{aligned} \tag{A5}$$

The function $u\ln u$ is negative on the whole interval $(0,\lambda/D)$, since $\lambda/D<1$. If we now send the cutoff Λ to infinity, the term with the integral in Eq. (A5) will go to zero as $1/\ln\Lambda$. The term $(\lambda/D)^{2mc\tau\lambda\Lambda}$ also vanishes, yielding the free-particle result

$$G(\tau,0) = -\ln\tau + \text{const.} \tag{A6}$$

APPENDIX B: CHERN-SIMONS ACTION FOR BULK CURRENTS

In this appendix we show that Eq. (4.8) is equivalent to the bulk action

$$S[A,g^i] = \frac{i\beta}{4\pi}\sum_{i=1}^{m}\int d^2x\,\varepsilon^{\mu\nu\kappa}[-(g_\mu^i)^\dagger\partial_\nu g_\kappa^i + 2(g_\mu^i)^\dagger\partial_\nu A_\kappa], \tag{B1}$$

with the condition $g_-^i = 0$ on the edge. The g^i's are $2+1$-dimensional potentials from which the electron current density j for every Landau level can be found:

$$j_i^\mu \propto \varepsilon^{\mu\nu\lambda}\partial_\nu g_\lambda^i. \tag{B2}$$

Notice three important subtleties.

(i) The coupling of g with the electromagnetic gauge field is of the form $\varepsilon^{\mu\nu\kappa}g_\mu\partial_\nu A_\kappa$ instead of the expected

$\varepsilon^{\mu\nu\kappa}A_\mu\partial_\nu g_\kappa \propto j_\mu A^\mu$. These expressions differ by an edge term. The second form is *not* invariant under the gauge transformations $A_\mu \to A_\mu + \partial_\mu\chi$; the expression $\varepsilon^{\mu\nu\kappa}\partial_\nu A_\kappa$, on the other hand, is manifestly gauge invariant.

(ii) Putting an arbitrary space-time component of g zero on the edge ensures that the action is invariant under $g_\mu \to g_\mu + \partial_\mu\kappa$, a gauge transformation that does not affect the current density. Without such a condition, gauge invariance is broken at the edge.

(iii) Because of the invariance under $g_\mu \to g_\mu + \partial_\mu\kappa$, a gauge fixing condition has to be specified for the path integration over g, for instance the Coulomb gauge $\nabla\cdot\vec{g}=0$.

Let us now for simplicity drop the replica indices α and the Landau level index i (effectively setting $m=1$). Having taken the condition $g_-|_{\text{edge}}=0$, the component g_- in Eq. (B1) multiplies the following constraint:

$$\nabla\times(\vec{g}-\vec{A})=0. \tag{B3}$$

After integration over g_-, what remains of the action is

$$\frac{i}{4\pi}\int d\tau\int d^2x(-\vec{g}\times\partial_-\vec{g}+2\vec{g}\times[\nabla A_- - \partial_-\vec{A}]), \tag{B4}$$

subject to constraint (B3). The general solution of Eq. (B3) is given by

$$\vec{g}=\vec{A}-\nabla\varphi, \tag{B5}$$

with $\varphi(\vec{x})$ a real scalar field which is now the only integration variable that is left. Substitution into Eq. (B4) yields an action where φ features only on the edge:

$$S[\varphi,A]=\frac{i}{4\pi}\int d\tau\Big[\int d^2x\,\varepsilon^{\mu\nu\kappa}A_\mu\partial_\nu A_\kappa - \oint dx(D_x\varphi D_-\varphi-\varphi E_x)\Big]. \tag{B6}$$

This is exactly the form of Eq. (4.8).

One may worry that the path integration over φ is ill defined, because of the bulk degrees of freedom of φ, which do not appear in Eq. (B6). However, φ inherits something from the gauge fixing condition of g. This is most easily seen in the case of the Coulomb gauge; here φ has to satisfy $\nabla^2\varphi=0$. This means that the bulk degrees of freedom are completely determined by $\varphi(x)$ at the edge (the well-known case of Laplace's equation with Dirichlet boundary conditions) and therefore are not independent integration variables.

One final remark on the boundary condition $g_-=0$: The Hamiltonian (density) corresponding to Eq. (B6) is given by $v_d(D_x\varphi)^2$. It is not allowed to choose a velocity $v_d<0$, since this would lead to energies that are unbounded from below. In general, the boundary condition has to be taken in such a way that the velocity of the chiral bosons has the same sign as the prefactor multiplying $i/4\pi$ in Eq. (B1), otherwise the integration over g is ill-defined on the edge.

APPENDIX C: INTERCHANNEL SCATTERING AT THE EDGE

In this appendix we describe the various steps of the standard Q-field approach to (edge) disorder. For the general case of m chiral edge channels, one can differentiate between different types of disorder, depending on whether one allows interchannel scattering or not. Although the different scattering potentials do not give rise to fundamentally different physical results, it is nevertheless important to define the ''effective'' edge Hamiltonian [Eq. (2.10)] which gives rise to the same result [Eq. (2.7)] that was previously obtained for 2D electrons. Below we shall show that the following m channel model satisfies our requirements

$$\mathcal{H}^{kk'}_{\text{edge}}=-iv_d\delta_{kk'}\partial_x+V_{kk'}(x), \tag{C1}$$

where V is a Hermitian random matrix and the elements $V_{kk'}$ are distributed with a Gaussian weight

$$P[V]=\exp\Big\{-\frac{1}{g}\oint dx\,\mathrm{tr}\,V^2\Big\}. \tag{C2}$$

The indices $k,k'=1,\ldots,m$ label the edge channels. The form (C1) implies that single potential scattering, as described by the 2D Hamiltonian

$$\mathcal{H}_{\text{2D}}=\frac{1}{2m_e}(\vec{p}-\vec{A})^2+V(\vec{x}), \tag{C3}$$

does not naively translate into single potential scattering for the edge states as obtained by solving Eq. (C3) in the presence of an edge (infinite potential wall). Rather than that, one should allow for interchannel scattering of the ''pure'' eigenstates as in Eq. (C1) in order to reproduce the effect of dirt in the general 2D problem (C3). We start from the following generating function for the averaged free-particle propagators:

$$Z=\int\mathcal{D}[\bar{\psi}\psi]\int\mathcal{D}[V]P[V]\times\exp\Big\{\beta\sum_{p=\pm,\alpha,jj'}\oint dx\,\bar{\psi}^{\alpha,j}_p\times[(\mu+ip\omega)\delta_{jj'}-\mathcal{H}^{jj'}_{\text{edge}}]\psi^{\alpha,j'}_p\Big\}. \tag{C4}$$

Integration over randomness and introduction of the matrix field $\tilde{Q}^{\alpha\beta}_{pp'}(x)$ by performing the Hubbard-Stratonovich trick leads to

$$Z=\int\mathcal{D}[\tilde{Q}]\exp\Big\{-\frac{1}{g}\mathrm{Tr}\,\tilde{Q}^2+m\,\mathrm{Tr}\,\ln[\mu+iv_d\partial_x+i\tilde{Q}+i\omega\Lambda]\Big\}. \tag{C5}$$

Notice that the edge channel label is not present in the field variable $\tilde{Q}$, but it is simply contained in an overall factor m. Notice also that the type of randomness as considered here has previously been introduced in a different context by the

name of N-orbital scattering, where N (here m) is commonly used for saddle-point and large-N expansion purposes.

We will next make use of the simple analytic properties of our 1D Hamiltonian and show that the saddle-point technique yields, in fact, exact results for all m and that therefore there is no need to rely on m to be ''large.'' The stationary point equation for $\tilde{Q}$,

$$i[\tilde{Q}_{\rm sp}]^{\alpha\beta}_{pp'}=\delta^{\alpha\beta}\delta_{pp'}[e_0+(-1)^p i/2\tau], \tag{C6}$$

can be written as

$$\frac{2}{g}(e_0\pm i/2\tau)=-m\int_{-\infty}^{\infty}\frac{dq}{2\pi}[\mu-v_d q+e_0\pm i(1/2\tau+\omega)]^{-1}$$

$$=\pm im/2v_d, \tag{C7}$$

with the simple solution $e_0=0$, $\tau=2v_d/(mg)$. One may next replace the original $\tilde{Q}$ field by the following change of variables:

$$\tilde{Q}\to T^{-1}PT\to\frac{1}{2\tau}T^{-1}\Lambda T=\frac{1}{2\tau}Q. \tag{C8}$$

Here $T\in {\rm SU}(2N)/{\rm S[U}(N)\times{\rm U}(N)]$ are unitary rotations and the block-diagonal Hermitian $P^{\alpha\beta}_{pp'}=\delta_{pp'}P^{\alpha\beta}_{p}$ represent the longitudinal components. Replacing P by its saddle-point value, as written in Eq. (C8), turns out to be an exact statement, valid for all m. The reason is contained in the fact that the fluctuations in P are weighted by propagators with poles in either the positive or negative imaginary momentum plane. All the momentum integrals therefore sum up to zero, giving rise to a zero weight to all orders in the P fluctuations. The replacement of Eq. (C8) is exact when inserted in the Tr ln. Equation (C5) factorizes into

$$Z=Z_P Z_T,$$

$$Z_P=\int\mathcal{D}[P]I[P]\exp\left\{-\frac{1}{g}{\rm Tr}\,P^2\right\}, \tag{C9}$$

$$Z_T=\int\mathcal{D}[T]\exp\left\{m\ {\rm Tr}\ \ln\left[\mu+iv_d\partial_x+\frac{i}{2\tau}\Lambda+iB\right]\right\}$$

where all T dependence is contained in the quantity B according to

$$B=v_d T\partial_x T^{-1}+\omega T\Lambda T^{-1}=v_d T D_0 T^{-1}. \tag{C10}$$

Equation (C9) can be evaluated further, and to lowest few orders in an expansion in B we obtain an effective action which can be written as

$$Z_T=\int\mathcal{D}[T]\exp S_{\rm eff}[T]$$

$$S_{\rm eff}[T]=\frac{m}{2v_d}{\rm Tr}\,\Lambda B(x)-\frac{m\tau}{8v_d}{\rm Tr}[B(x),\Lambda]^2+\cdots \tag{C11}$$

$$=\frac{m}{2}\oint dx\ {\rm tr}\,\Lambda T\partial_x T^{-1}+\frac{m}{2v_d}\omega\oint dx\ {\rm tr}\,\Lambda Q$$

$$-\frac{m\tau v_{\rm d}}{8}\oint dx\ {\rm tr}[D_0,Q]^2. \tag{C12}$$

The coefficients appearing in Eq. (C12) all have a clear physical significance in the context of disordered edge states (see also the main text). In particular, m stands for the quantized Hall conductance σ_{xy}; $m/2\pi v_{\rm d}$ equals the total density of edge states $\rho_{\rm edge}$. The quantity $m\tau v_{\rm d}$ that appears in the higher-dimensional operators is the 1D conductivity σ_{xx} of m channel edge states. Here $2\tau v_d$ is the linear dimension which sets the smallest wavelength for the Q field variables, and $m/2$ is the (quantized) conductance (g_m) of the wire.

APPENDIX D: ACTION FOR Q AND A ON MULTIPLE EDGES

The generalization of Eq. (4.11) is given by

$$S[Q,A]=\frac{\beta/2}{(2\pi)^2}\int d^2x\,d^2x'\,n(\vec{x})B^\dagger(\vec{x})U_0(\vec{x}-\vec{x}')n(\vec{x}')B(\vec{x}')$$

$$+\frac{i\beta}{4\pi}\left[\int d^2x\,n(\vec{x})\varepsilon^{\mu\nu\kappa}(A^{\rm eff}_\mu)^\dagger\partial_\nu A^{\rm eff}_\kappa+\sum_{a=1}^{M}s_a\oint_{C_a}dx\left(A^\dagger_x A^{\rm eff}_0-\frac{2\pi}{\beta}{\rm tr}\,\hat{A}_x Q\right)\right]$$

$$+\sum_{a=1}^{M}s_a S^{(a)}_{\rm top}[Q]+\frac{\pi}{4\beta v_d}\sum_{a=1}^{M}S^{(a)}_F[Q]-\frac{\pi}{4\beta}\sum_{n\alpha}\sum_{a=1}^{M}\int_{C_a}\frac{dk_x}{2\pi}\frac{1}{v^{\rm eff}(k_x)}\left|{\rm tr}\,I^\alpha_n Q(k_x)-\frac{\beta}{\pi}(A^{\rm eff}_0)^\alpha_{-n}(k_x)\right|^2 \tag{D1}$$

$$+\frac{1}{8\beta v_d^2}\sum_{a\neq b}s_a s_b\oint_{C_a}dx\oint_{C_b}dx'\sum_{n\alpha}\left[{\rm tr}\,I^\alpha_n Q-\frac{\beta}{\pi}(A^{\rm eff}_0)^\alpha_{-n}\right](x)$$

$$\times U_0(x,x')\left[{\rm tr}\,I^\alpha_{-n}Q-\frac{\beta}{\pi}(A^{\rm eff}_0)^\alpha_n\right](x'), \tag{D2}$$

where $U_0(x,x')$ denotes the full 2D Coulomb interaction. All terms except those quadratic in Q arise by the obvious replacements $m \to n(\vec{x})$ and $m \oint \to \Sigma_a s_a \oint_{C_a}$ in Eq. (4.11). The terms quadratic in Q can be understood as follows. In the generalized form of Eq. (3.39), the quadratic term in the plasmon field is given by

$$-\frac{\beta}{2}\int d^2x\, d^2x'\, \lambda(x)^\dagger U_0^{-1}(x-x')\lambda(x') + \frac{m}{2\pi v_d}\sum_a \oint_{C_a} dx\, \lambda^\dagger\lambda, \tag{D3}$$

indicating that the propagator for λ between two points on the same edge will be very different from the propagator between different edges. In the former case the propagator is proportional to $[U_0^{-1}+(m/2\pi v_d)]^{-1}$, which is exactly the form obtained by combining the Finkelstein term with Eq. (D1). In the latter case, the propagator is simply proportional to U_0. Finally, the signs can be understood by noticing that the coupling of the plasmon field to Q is proportional to $\Sigma_a s_a \oint_{C_a} dx\, \mathrm{tr}\,\hat{\lambda} Q$.

*Permanent address: Russian Research Center ''Kurchatov Institute,'' Kurchatov sq. 1, 123182 Moscow, Russia.

[1] X. G. Wen, Phys. Rev. B **41**, 12 838 (1990); Int. J. Mod. Phys. B **6**, 1711 (1991); cond-mat/9506066 (unpublished).

[2] J. Fröhlich and A. Zee, Nucl. Phys. B **364**, 517 (1991).

[3] C. L. Kane and M. P. A. Fisher, Phys. Rev. B **51**, 13 449 (1995); C. L. Kane, M. P. A. Fisher, and J. Polchinski, Phys. Rev. Lett. **72**, 4129 (1994).

[4] F. D. M. Haldane, Phys. Rev. Lett. **74**, 2090 (1994).

[5] M. Grayson, D. C. Tsui, L. N. Pfeiffer, K. W. West, and A. M. Chang, Phys. Rev. Lett. **80**, 1062 (1998).

[6] A. M. M. Pruisken, M. A. Baranov, and B. Škorić, cond-mat/9712322 (unpublished).

[7] M. A. Baranov, A. M. M. Pruisken, and B. Škorić, cond-mat/9712323 (unpublished).

[8] A. M. M. Pruisken, Nucl. Phys. B **235**, 277 (1984); **285**, 719 (1987); **290**, 61 (1987).

[9] A. P. Balachandran, L. Chandar, and B. Sathiapalan, Nucl. Phys. B **443**, 465 (1995).

[10] R. E. Prange and S. M. Girvin, *The Quantum Hall Effect* (Springer, Berlin, 1990).

[11] See A. M. M. Pruisken, in *The Quantum Hall Effect* (Ref. 10).

[12] A. M. M. Pruisken and H. P. Wei, *Ordering Disorder: Prospect and Retrospect in Condensed Matter Physics*, edited by Vipin Srivastava, Anil K. Bhatnagar, and Donald G. Naugle, AIP Conf. Proc. No. 286 (AIP, New York, 1993), p, 215.

[13] D. Shahar *et al.*, Solid State Commun. **107**, 19 (1998) N. Balaban, U. Meirav, and I. Bar-Joseph, cond-mat/9806214 (unpublished).

[14] B. Škorić and A. M. M. Pruisken, Nucl. Phys. B: Field Theory Stat. Syst. **599**[FS], 637 (1999).

[15] A. M. M. Pruisken, Phys. Rev. Lett. **61**, 1298 (1988).

[16] H. P. Wei, D. C. Tsui, M. A. Paalanen, and A. M. M. Pruisken, Phys. Rev. Lett. **61**, 1294 (1988).

[17] B. Huckestein and B. Kramer, Phys. Rev. Lett. **64**, 1432 (1990).

[18] A. M. M. Pruisken and M. A. Baranov, Europhys. Lett. **31**, 543 (1995).

[19] A. M. M. Pruisken and Z. Wang, Nucl. Phys. B **322**, 721 (1989).

[20] A. M. M. Pruisken, Phys. Rev. B **31**, 416 (1985).

[21] See R. E. Prange, in *The Quantum Hall Effect* (Ref. 10).

[22] J. T. Chalker and P. D. Coddington, J. Phys. C **21**, 2665 (1988).

[23] H. Levine, S. B. Libby, and A. M. M. Pruisken, Nucl. Phys. B **240**, 30 (1984); **240**, 49 (1984); **240**, 71 (1984).

[24] E. Witten, Nucl. Phys. B **149**, 285 (1979).

[25] A. M. M. Pruisken, M. A. Baranov, and B. Škorić (unpublished).

[26] We frequently use the † notation, introduced in Paper I, in order to simplify the replica and Matsubara frequency representation of bilinear forms. This notation is often extended to include the familiar space-time symbols. For example, the three-vector $z^\dagger \partial_\mu z$ actually stands for

$$z^\dagger \partial_0 z = \sum_{\alpha n} z^\alpha_{-n}(-i\nu_n) z^\alpha_n, \quad z^\dagger \partial_i z = \sum_{\alpha n} z^\alpha_{-n} \partial_i z^\alpha_n,$$

where ν_n is the bosonic Matsubara frequency and α, n denote the replica and frequency indices, respectively.

[27] Y. Oreg and A. M. Finkelstein, Phys. Rev. Lett. **74**, 3668 (1995).