

Duality near quantum Hall transitions

E. Shimshoni

Department of Mathematics-Physics, Oranim-Haifa University, Tivon 36006, Israel

S. L. Sondhi

Department of Physics, Princeton University, Princeton, New Jersey 08544

D. Shahar

Department of Electrical Engineering, Princeton University, Princeton, New Jersey 08544

(Received 9 October 1996)

A recent experiment by Shahar *et al.*, on the phase transitions between quantum Hall states and the insulator, found that the current-voltage characteristics in the two phases are related by symmetry. It was suggested in this work that this is evidence for charge-flux duality near quantum Hall transitions. Here we provide details of this analysis. We review some theoretical ideas on charge-flux duality in the composite boson description of the quantum Hall effect, and interpret the data as implying that this duality is a symmetry in the transition region *and* that the Hall response of the bosons vanishes. We observe that duality for composite bosons is equivalent to a particle-hole symmetry for composite fermions and show that a Landauer analysis of transport for the latter allows a possible understanding of the reflection symmetry and Hall response beyond the linear regime. We note that the duality interpretation supports the scenario of superuniversality for quantum Hall transitions outlined by Kivelson, Lee, and Zhang. Finally, we discuss how to search for the duality at other transitions in the quantum Hall regime. [S0163-1829(97)00420-7]

I. INTRODUCTION

Two-dimensional electron systems in the quantum Hall (QH) regime, exhibit an extremely complex phase diagram with a rich set of transitions between various phases. In recent years, a combination of theoretical and experimental work has led to considerable progress in understanding this complexity. An important milestone was the work of Kivelson, Lee, and Zhang² (KLZ) who proposed a topology for the phase diagram (in the disorder-magnetic field plane) on the basis of a set of correspondence rules. These codified two ideas which are central to the current understanding of the quantum Hall effect (QHE). The first of these is the existence of sets of QH states that are related by flux attachment transformations that are the basis of Chern-Simons theories of the QHE (Ref. 3) and of the work of Jain.⁴ The second idea is the hierarchical structure of the set of QH states, i.e., that, starting with a basis of primary (Laughlin) states, one can construct members of the set by “condensing” quasiparticles of other members of the set.⁵ In the bosonic Chern-Simons framework used by KLZ, quasiparticles are vortices and hence the hierarchical descendants are obtained as saddle points of actions obtained by repeated duality transformations. The topology of the phase diagram is then fixed by requiring that neighboring phases be related by the condensation of *one* set of quasiparticles, or by one duality transformation. Experimentally, this expectation has been borne out fairly well and although there is evidence of systematic departures at low fields and in the reentrant region near $\nu=1/5$,⁶ these are not an issue in the region of interest to us in this paper.

The transitions between the various phases in the QH re-

gime have been the focus of another active line of work, particularly following the contributions of Pruisken⁷ and Wei, Tsui, Paalanen, and Pruisken⁸ who proposed that the transitions between integer QH states are quantum phase transitions, i.e., continuous $T=0$ phase transitions. Subsequent experimental work has supported this identification for transitions between other QH states and has suggested a hypothesis of “super-universality,” i.e., that *all* such transitions share the same exponents, and a tentative identification of the correlation length exponent ν_ξ and the dynamic scaling exponent z .^{9,10} remarkably, the value of ν_ξ agrees with numerical calculations for noninteracting electrons in the lowest Landau level.¹¹ In the work of KLZ, an explanation was offered for the superuniversality (along with an extension to the QH/insulator transition of interest to us) in terms of a Gaussian analysis of the fluctuations of the Chern-Simons field which allowed them to predict a set of critical conductivities for the various transitions as well. While the latter predictions have gathered some experimental support, their neglect of higher order fluctuations has been criticized and defended in model calculations.^{12,13} In Sec. IV we comment on the issues in this debate.

A recent experiment of Shahar *et al.*,¹ on the transition between the $1/3$ QH fluid and the proximate, high field insulator, appears to tie these two streams of ideas together in a very interesting manner while also lending strong support to the account of KLZ of the critical behavior. To recapitulate, the experiment found that for a range of fields near the transition between the QH state and the insulator, the transport data in the two phases are related by an unanticipated symmetry. More precisely, there exist filling fractions (ν, ν_d) in the QH and insulating states at which the *nonlinear* longitu-

dinal current density/electric field (j, E_L) characteristics are related by *reflection*:

$$\{j(\nu_d), E_L(\nu_d)\} = \left\{ \frac{e^2}{h} E_L(\nu), \frac{h}{e^2} j(\nu) \right\}. \quad (1)$$

On the basis of the functional relationship between ν_d and ν , this was identified in Ref. 1 as a manifestation of charge-flux duality for composite bosons and equivalently, particle-hole symmetry for composite fermions. This was then used to argue that the Hall response must also be symmetric and have the particularly simple form of the (j, E_H) characteristics being *linear* and unchanged across the transition,

$$E_H = 3 \frac{h}{e^2} j \nu, \quad (2)$$

an expectation borne out by the data presented in Ref. 1. Similar results hold at the $\nu=1$ to insulator transition,¹⁴ implying that the symmetry is present at the transitions between primary QH fluids and the insulator. [In the rest of the paper, we shall use *reflection symmetry* as a shorthand for the experimentally observed symmetry of the response summarized in Eqs. (1) and (2) while reserving the term duality for the theoretical inference about the underlying dynamics.]

In this paper we present the theoretical underpinnings of the interpretation advanced in Ref. 1. In Sec. II we describe these in the language of composite bosons with the help of the Chern-Simons resistivity addition law and duality transformations. We comment there on the issue of whether the critical action is Gaussian. Next, in Sec. III we introduce a Landauer formulation of the transport in the composite fermion basis that connects the experimental observations to particle-hole symmetry for the fermions—which is shown to be equivalent to duality for the bosons. In both cases we note that an additional postulate, equivalent to the constancy of the Hall response, needs to be introduced separately.

In Sec. IV we switch gears, and adduce purely theoretical evidence that this symmetry is present at long wavelengths by revisiting existing studies of QH transitions. In Sec. V, we discuss how the symmetry is hidden at interplateau transitions and show how it can be detected by a deconvolution of the experimental data. We end with a summary of our argument and some key open issues in our approach. As our work draws on much previous work in addition to that reviewed already, most notably that of Fisher¹⁵ and of Ruzin and collaborators,^{16,17} we comment on these connections in the main text.

II. THE BOSONIC APPROACH—DUALITY

A. Dual filling factors

We begin by recapitulating the composite boson interpretation of the transition between the primary Hall states and the insulator. A Chern-Simons transformation^{2,18} is used to represent the physical electrons as bosons that carry k flux quanta (ϕ_0), k being an odd integer. At the filling factor ν , the resulting bosons experience a mean magnetic field $B_{\text{eff}} = B(1 - k\nu)$, which vanishes at $\nu = 1/k$ thus enabling them to condense into a superconducting state. A decrease in ν results in a nonzero B_{eff} which induces vortices. At first the vortices are trapped by disorder and the bosons continue to

superconduct. However at a critical density the vortices delocalize and destroy the superconductivity of the bosons. This causes the bosons to become localized and the original electronic system becomes insulating as well. The composite boson picture thus naturally associates the physics of QH/insulator transitions with that of field-tuned superconductor/insulator transitions.¹⁹

This description is clearly marked by a *duality* in the roles of the bosons/vortices which are localized in the insulating/superconducting phases, respectively, and are delocalized in the conjugate phase. This suggests that the problem might be characterized by a *symmetry*, which we shall also refer to as duality,²⁰ that relates the superconducting and insulating phases in that the dynamics of the system expressed in terms of the bosons at a filling ν in the QH phase is identical to the dynamics at ν_d in the insulating phase expressed in terms of the vortices. The natural candidates for the dual fillings are related by

$$\frac{n_v(\nu)}{n_b(\nu)} = \frac{n_b(\nu_d)}{n_v(\nu_d)}, \quad (3)$$

where n_b and n_v are the boson and vortex densities, respectively. As the vortex density is set by the effective field, $\rho_v = B_{\text{eff}}/\phi_0 = \rho_b(\nu^{-1} - k)$, it follows that

$$\frac{1}{\nu^{-1} - k} = \nu_d^{-1} - k. \quad (4)$$

Explicitly, $\nu_d = (1 - k\nu)/[\nu + k(1 - k\nu)]$.¹ The observation that the filling fractions related by the reflection symmetry in the experiment obey Eq. (4) to a good approximation, strongly suggests that the reflection symmetry is a manifestation of duality.

To further this interpretation, we need to examine whether the reflection symmetry follows from an assumption of identity between the boson and vortex actions at ν and ν_d , respectively. In order to do this we rederive the well known Chern-Simons resistivity relation²¹ that relates the bosonic response to the electronic response, in a way that suggests its validity beyond the linear regime.

B. Chern-Simons resistivity addition

Our starting point is the bosonic Chern-Simons functional integral¹⁸ for the ‘‘Helmholtz’’ functional $F[A_\mu; \nu]$,²² of the electrons at filling fraction $0 \leq \nu \leq 1/k$, in the presence of an external three-vector gauge field A_μ ($\hbar = e/c = 1$),

$$e^{-F[A_\mu; \nu]} = \int D[a_\mu] e^{-(i/2k)S_{\text{CS}}} e^{-F^b[A_\mu + a_\mu; \nu]},$$

$$S_{\text{CS}} \equiv \int d\mathbf{r} d\tau \epsilon_{\mu\nu\lambda} a_\mu \partial_\nu a_\lambda. \quad (5)$$

Here,

$$e^{-F^b[A_\mu + a_\mu; \nu]} = \int D[\phi] e^{-S_M(\phi, A_\mu + a_\mu; \nu)} \quad (6)$$

defines the free energy of the composite bosons, which is obtained after (exactly) integrating out the fluctuations of the boson field. The observable electromagnetic response (asso-

ciated with functional derivatives of F) is thus related to the response of the bosons to the effective field $A_\mu + a_\mu$. In linear response, the free energies are quadratic in the gauge fields, with coefficients that are directly related to the conductivity tensor computed in the random phase approximation (RPA). The integration in Eq. (5) can be performed explicitly, resulting in an algebraic relation between $\sigma_{ij}, \sigma_{ij}^b$ of the physical electrons and bosons, respectively. This relation is most simply phrased in terms of the *resistivity* tensors. Here we proceed slightly differently, in a manner which might generalize to nonlinear response.

Our strategy is to formally introduce the resistivity tensor before carrying out the integration over a_μ in Eq. (5). We convert from the Helmholtz functional $F[\mathbf{A}; \nu]$ to the ‘‘Gibbs’’²² functional $G[\mathbf{J}; \nu]$ via a functional Fourier transform:

$$e^{-G[\mathbf{J}; \nu]} = \int D[\mathbf{A}] e^{-F[\mathbf{A}; \nu]} e^{-i\int \mathbf{A} \cdot \mathbf{J}}. \quad (7)$$

(For quadratic actions, this can substitute for the standard Legendre transform and therefore allows a purely functional integral argument. It is unclear whether something of this kind can be done beyond quadratic order—this will be further discussed elsewhere.²³) Hereon we choose the gauge $A_0 = 0$ for all vector potentials, so that the corresponding electric fields are (in Fourier space) $E_j = i\omega A_j$. To quadratic order, the functionals F and G are given by

$$F[\mathbf{A}; \nu] = \int d^3q \frac{1}{2} \sum_{i,j} \omega A_i(q) \sigma_{ij}(q; \nu) A_j(-q),$$

$$G[\mathbf{J}; \nu] = \int d^3q \frac{1}{2} \sum_{i,j} (1/\omega) J_i(q) \rho_{ij}(q; \nu) J_j(-q), \quad (8)$$

where $q = (\omega, \mathbf{q})$ and $i, j \in \{x, y\}$. In the same way, we define $G^b[\mathbf{J}]$ for the bosons. We then insert Eq. (5) into the definition of $G[\mathbf{J}]$, Eq. (7). Integrating over $A_\mu + a_\mu$, and using the corresponding definition of $G^b[\mathbf{J}]$, this yields

$$e^{-G[\mathbf{J}; \nu]} = e^{-G^b[\mathbf{J}; \nu]} \int D[\mathbf{a}] e^{iS[\mathbf{a}]},$$

$$S[\mathbf{a}] \equiv \int d^3q \left\{ \left(\frac{i\omega}{k} \right) a_x(-q) a_y(q) + \mathbf{a}(-q) \cdot \mathbf{J}(q) \right\}. \quad (9)$$

The integration over \mathbf{a} then yields

$$G[\mathbf{J}; \nu] = G^b[\mathbf{J}; \nu] + \int d^3q (k/\omega) J_x(q) J_y(-q). \quad (10)$$

It follows that the dc resistivities of the electrons and bosons are related by the addition of the Chern-Simons resistivity, i.e.,

$$\rho = \rho^b + \rho^{\text{cs}}(k), \quad (11)$$

where

$$\rho = \begin{pmatrix} \rho_L & \rho_H \\ -\rho_H & \rho_L \end{pmatrix},$$

$$\rho^b = \begin{pmatrix} \rho_L^b & \rho_H^b \\ -\rho_H^b & \rho_L^b \end{pmatrix},$$

$$\rho^{\text{cs}}(k) = \begin{pmatrix} 0 & k \\ -k & 0 \end{pmatrix}, \quad (12)$$

are resistivities scaled by h/e^2 . (We follow this convention hereafter.) This is frequently summarized by the statement that the measured voltage is the sum of the bosonic response and the Faraday effect stemming from the flux carried by them. This last interpretation is clearly not limited to linear response and hence we might expect Eq. (11) and some version of Eq. (10) to hold in the nonlinear region as well. (This extension can also be argued by proceeding from Ehrenfest’s theorem applied to the operator equations of motion for the Chern-Simons Lagrangian of Ref. 18. The real problem here, discussed further in Sec. III B, is the lack of a fluctuation-dissipation theorem beyond the linear regime.)

In proceeding from Eq. (5) to Eq. (11), we have ignored a number of delicate issues regarding disorder averages and the order of the k, ω , and T limits. On the latter point we have in mind the ordering $k \ll \omega \ll T$, so that we are in the correct regime for defining transport coefficients and mesoscopic fluctuations are not an issue on account of dephasing processes at $T \neq 0$. Consequently, our functional integrals are implicitly bounded between $\tau = 0$ and $\tau = \hbar\beta$. Giving meaning to the boson transport coefficients is more subtle. Outside of our formal manipulations with functional integrals, the best approach is to think of the conductivities as being given diagrammatically by the sum of graphs that are one-particle irreducible with respect to the Chern-Simons interaction but are permitted to have an arbitrary set of internal Chern-Simons, disorder, and scalar interaction lines. With this definition, Eq. (11) is a tautology.

C. Duality and transport: Phenomenology

We now return to our main argument and consider the implications of duality for ρ^b ; in the next section we will interpret the reflection symmetry in their light. Our discussion here is a generalization to the nonlinear case of Fisher’s discussion for the field-tuned superconductor insulator transition.¹⁵ By hypothesis, the boson and vortex dynamics are the same at dual fillings except that they see oppositely directed magnetic fields. Hence the boson and vortex resistivity matrices at these fillings are related by

$$\rho^b(j, \nu) = [\rho^v(j_v, \nu_d)]^\dagger, \quad (13)$$

where the explicit dependence on the boson and vortex currents indicates the possibly nonlinear nature of the response. As bosons and vortices see electric fields and currents oppositely, their resistivity and conductivity matrices at the *same* filling are related by

$$\rho^v(j_v, \nu_d) = \sigma^b(E, \nu_d). \quad (14)$$

Combining these we find

$$\rho^b(j, \nu) = [\sigma^b(E, \nu_d)]^\dagger, \quad (15)$$

which summarizes the implications of duality for the transport coefficients in and beyond the linear regime.

In the linear regime, Eq. (15) can be written explicitly as a pair of relations between the resistivities,

$$\begin{aligned}\rho_L^b(\nu) &= \frac{\rho_L^b(\nu_d)}{[\rho_L^b(\nu_d)]^2 + [\rho_H^b(\nu_d)]^2}, \\ \rho_H^b(\nu) &= \frac{\rho_H^b(\nu_d)}{[\rho_L^b(\nu_d)]^2 + [\rho_H^b(\nu_d)]^2}.\end{aligned}\quad (16)$$

We see that, in general, these relate either $\rho_L^b(\nu)$ or $\rho_H^b(\nu)$ to *both* $\rho_L^b(\nu_d)$ and $\rho_H^b(\nu_d)$, i.e., duality does not always preserve the longitudinal/Hall distinction. [At the self-dual critical point, $\nu_d = \nu$, Eq. (16) leads to the constraint $(\rho_L^b)^2 + (\rho_H^b)^2 = 1$ on the critical resistivities.]

D. Duality and transport: Experiments

Turning now to the experimental data, we note that Eqs. (1) and (2) imply that the longitudinal bosonic response obeys

$$\rho_L^b(\nu, j) = \sigma_L^b(\nu_d, E) \quad (17)$$

(j and E have the same numerical value in our units), and that the bosonic Hall response vanishes, i.e.,

$$\rho_H^b(\nu, j) = 0 \quad (18)$$

for all ν in the transition region. These evidently satisfy the constraint in Eq. (15) in a particularly simple way. It follows then that it is necessary but not sufficient to postulate duality in order to account for the data—but duality and a vanishing bosonic Hall response are both necessary and sufficient.

Let us briefly speculate on possible constraints on the form of the critical point action from the reflection symmetry. If Eq. (10) could be placed on a secure footing, the reflection symmetry would suggest that the bosonic Gibbs functional, defined as the Legendre transform of the bosonic Helmholtz functional, has the separable form,

$$G^b[\mathbf{J}] = G_l^b[J_x] + G_l^b[J_y], \quad (19)$$

with the duality constraint

$$G_l^b[J_j; \nu] = F_l^b\left[\frac{1}{i\omega} J_j; \nu_d\right]. \quad (20)$$

At the critical point this would require a functional form invariant under a Legendre transform, i.e., a Gaussian.

E. Duality and transport: Duality transformations

In our discussion thus far, we have relied upon the phenomenology of bosons and vortices familiar from the context of superfluidity/superconductivity and the resistivity addition law with a fixed Chern-Simons coefficient. Now we will present a derivation of Eqs. (16) from the viewpoint of a duality transformation in which we trade the bosons and Chern-Simons field for another set of bosons and a different Chern-Simons field. In doing so we will make a crucial assumption, which is less restrictive than the one made by KLZ in their original analysis but is very much in their spirit, which will thus be seen to be equivalent to the assumption of duality.

The formulation of the duality transformation that we will need is the following. The bosonic Chern-Simons path integral expression for the partition function of the electron gas at filling fraction ν ,

$$Z[A_\mu] = \int D[a_\mu] \int D[\phi] e^{-(i/2k)S_{\text{cs}}(a_\mu) - S_M(\phi, A_\mu + a_\mu; \nu)}, \quad (21)$$

is believed to be rewritable,²⁴ up to irrelevant terms, as another bosonic path integral that differs only in the Chern-Simons coefficient, the charge of the bosons and their filling fraction,

$$\begin{aligned}Z[A_\mu] &= e^{-(i/2k)S_{\text{cs}}(A_\mu)} \int D[\tilde{a}_\mu] \\ &\times \int D[\tilde{\phi}] e^{-i(k/2)S_{\text{cs}}(\tilde{a}_\mu) - S_M[\tilde{\phi}(1/k)A_\mu + \tilde{a}_\mu; \nu_d]}.\end{aligned}\quad (22)$$

Note that the transformed bosons, which are the vortices, conduct in parallel to the ideal QH fluid with $\sigma_{xy} = 1/k$.

It follows from Eq. (21) that the resistivity at ν takes the form

$$\rho(\nu) = \rho^b(\nu) + \rho^{\text{cs}}(k), \quad (23)$$

where ρ^b is the boson resistivity. Similarly it follows from Eq. (22) that the conductivity at ν_d can be written as

$$\sigma(\nu_d) = \sigma^{\text{cs}}(k) + \{k^2 \tilde{\rho}^b(\nu) + k^2 \rho^{\text{cs}}(-k)\}^{-1}, \quad (24)$$

where $\sigma^{\text{cs}}(k) = [\rho^{\text{cs}}(k)]^{-1}$ is the Chern-Simons conductivity and $\tilde{\rho}^b$ is the *dual* boson resistivity, i.e., the resistivity of the bosons in the dual representation. If we make the assumption that this is the *same* function of filling fraction as ρ^b we find that $\rho(\nu_d)$ can be rewritten in the form of Eq. (23) with $\rho^b(\nu_d) = [\sigma^b(\nu)]^\dagger$ as before. We emphasize that this is not a trivial assumption—the two sets of bosons interact with gauge fields governed by Chern-Simons terms with different coefficients and therefore their resistivities might be expected to exhibit different functional dependences on ν . Of course, the requirement that the $\rho(\nu)$ be the same in both representations does relate $\rho^b(\nu)$ with $\tilde{\rho}^b(\nu_d)$.

III. THE FERMIONIC APPROACH—PARTICLE-HOLE SYMMETRY

A. Symmetric fillings and equivalence with duality

We will now consider the QH/insulator transitions in the composite fermion description.²⁵ In this, the electronic state at ν is related by an even flux attachment (Chern-Simons) transformation to a state of composite fermions at the auxiliary filling fraction $\nu' = \nu/[1 - (k-1)\nu]$; here $k-1$, with k odd, is the number of flux quanta. This maps $\nu=1/k$ to $\nu'=1$ and the transition to the insulator has the form of the depletion of a single filled (pseudo) Landau level, i.e., the $\nu=1 \rightarrow 0$ transition. (However, the transitions are *prima facie* different, as the $k \neq 1$ transitions involve gauge fields not present in the $k=1$ case.)

As the space of states involved in the transition now includes a full Landau level, it becomes possible to formulate a

particle-hole symmetry for the fermions that asserts that the dynamics of the particles at auxiliary filling ν' is the same as that of holes at auxiliary filling $\nu'_d = 1 - \nu'$. In fact, this symmetry is the same as duality for the composite bosons as can be seen by rewriting the particle-hole condition in terms of the filling factor itself and noting its equivalence with the duality condition derived earlier. Again, without addressing the microscopic origin of the symmetry, we will explore its consequences for the electronic transport. In the following we will argue that it implies the reflection symmetry, provided that we accept a particular Landauer framework for the analysis of the transport.

B. Particle-hole symmetry and transport

Much as in the bosonic description, transport coefficients for the electrons and the fermions are related by another addition relation,

$$\rho = \rho^f(\nu) + \rho^{cs}(k-1), \quad (25)$$

where ρ^f is the fermion resistivity. Before turning to the implications of particle-hole symmetry, we will sketch a Landauer theory for ρ^f along the lines of Jain and Kivelson's treatment of noninteracting electrons.²⁶ Our motivation here is twofold. First, we take the experiments to date, with their surprising estimate of a correlation length exponent consistent with noninteracting calculations, as suggesting that a composite fermion quasiparticle description continues to hold even in the transition regions so that we can use an effective single particle description for them. Next, we feel that if the experiments are measuring universal transport data in a critical region, it should not matter exactly how we compute these quantities; i.e., the universal part of the transport might be computable by our idealized Landauer calculation even if there are nonuniversal parts sensitive to the actual arrangements of contacts in the device. Needless to say, this assumption needs further study.²⁷

With these caveats, we consider the conductances of a disordered QH region sandwiched between two ideal regions that serve to define incoming and outgoing edge states with linear dispersions.²⁸ We imagine a calculation in the critical region where the size of the disordered region is set by a dephasing length. The transport through the region is characterized by an energy dependent transmission coefficient $\mathcal{T}(\epsilon)$. Following Jain and Kivelson, we define the current I , longitudinal voltage V_L , and Hall voltage V_H for given edge chemical potentials $\mu_1 = \epsilon_F - V/2$ and $\mu_2 = \epsilon_F + V/2$ (ϵ_F is the equilibrium chemical potential on the both edges) as

$$\begin{aligned} I &= \int_{\mu_L}^{\mu_R} d\epsilon \mathcal{T}(\epsilon), \\ V_L &= (\mu_2 - \mu_1) - \int_{\mu_L}^{\mu_R} d\epsilon \mathcal{T}(\epsilon), \\ V_H &= \int_{\mu_1}^{\mu_2} d\epsilon \mathcal{T}(\epsilon). \end{aligned} \quad (26)$$

Evidently, these expressions allow for nonlinear dependences of the currents and voltages on the source-drain voltage $V = \mu_2 - \mu_1$ and hence on each other; all that this re-

quired is for $\mathcal{T}(\epsilon)$ to have structure on the scale of V . In the transition region, this is easily arranged for the variation of $\mathcal{T}(\epsilon)$ becomes arbitrarily rapid at low temperatures (it is a step function at zero temperature). Consequently, one should generically expect nonlinear transport from the proximity to the critical point.

Nevertheless, the most striking feature of Eqs. (26) is the linear dependence of V_H on I independent of ϵ_F and hence of ν . This striking feature of the experimental data follows automatically from Jain and Kivelson's Landauer formalism. If, in addition, we postulate particle-hole symmetry ($\epsilon_c = 0$),

$$\mathcal{T}(\epsilon) = 1 - \mathcal{T}(-\epsilon), \quad (27)$$

we find that in going between symmetry related fillings, $\epsilon_F \rightarrow -\epsilon_F$, V_L and I simply trade places. Such a treatment can then account naturally for the symmetry even in the nonlinear transport.

However, an important caveat is in order. The Landauer formalism assumes that the dissipation necessary to produce steady state transport takes place in the reservoirs defining the edge chemical potentials and hence the nonlinearities it produces are due to elastic physics alone. In a real system one needs to worry about nonlinearities arising from dissipative bottlenecks, possibly coming from critical physics themselves.²⁹ In other words, there is no fluctuation-dissipation theorem for nonlinear response which relates it uniquely to equilibrium correlations and, in principle, it is necessary to include explicit dissipative mechanisms in calculations.

Evidently, we have not done so in our treatment of the nonlinear response in the bosonic description and have chosen a particular, infinitely efficient, mechanism in the fermionic Landauer description. For our purposes, the importance of the latter is that it is a proof of principle that intrinsic critical physics *can* lead to the nonlinear symmetry observed in the data.

IV. THEORETICAL EVIDENCE FOR DUALITY/ PARTICLE-HOLE SYMMETRY AND A CONSTANT ρ_H

Our discussion so far has focused on phenomenology in that we have attempted to translate the experimental observations into the framework of the Chern-Simons description of the QH transitions. This has led us to conclude that the composite boson description must be marked by duality and a vanishing Hall response and the composite fermion description by particle-hole symmetry and a constant Hall response. In this section we will review previous theoretical work which, though it did not anticipate the particular striking features of the data, does suggest that our inferences would arise naturally in a microscopic theory of the QH phase transitions.

Duality. In Sec. II E we noted the belief, based on the work of Lee and Fisher,²⁴ that the bosonic Chern-Simons action with coefficient k for particles of charge 1 at filling fraction ν , could be written as another, dual Chern-Simons action with coefficient $1/k$ for particles of charge $1/k$ at filling fraction ν_d . The caveats necessary here are (a) that their arguments on the irrelevance of other terms generated in the dual action are compelling deep in the QH phases, but do not

take account of any anomalous dimensions that would be produced near a $T=0$ critical point, and (b) that the introduction of disorder should be expected to change the precise functional connection between ν and ν_d . We also noted that for duality to be a symmetry in the sense of Eqs. (16), it is necessary that the dual actions lead to the same $\rho^b(\nu)$.

The latter assumption was made by KLZ, formally based on a problematic RPA treatment of the disorder which ignores any internal Chern-Simons lines in the diagrams that contribute to ρ^b . This treatment was subsequently called into question by calculations on undistorted systems that suggested that the Chern-Simons term is generically a marginal perturbation and gives rise of a line of fixed points with continuously varying exponents;¹² diagrammatically, they showed that the neglect of internal Chern-Simons lines is not always justified.

Nevertheless, there is a second, qualitative argument in support of the claim by KLZ that appears more robust and has been the inspiration for some recent, competing, model calculations for the defense.¹³ If we accept the long wavelength form invariance of the original and dual actions, then it follows that universal quantities computed from them, internal gauge field lines and all, can depend only upon the filling fraction (or the appropriate scaling field), the charge of the bosons, and the Chern-Simons coefficient. As both actions describe the same transition, it follows that they must yield the same correlation length/time exponents despite the differing charge and statistics. That suggests, though it does not dictate, that they yield the same scaling functions for the resistivities; indeed, for particle-hole asymmetric disorder the converse would seem to be a serious possibility (see below). At any rate it calls into question the relevance of calculations where the exponents *do* vary with the Chern-Simons coupling. (These calculations¹² perturb in the Chern-Simons coefficient and hence are inconclusive on the large coefficient shift $k \rightarrow 1/k$ at issue in the question of duality. Recent model calculations¹³ have attempted to show that models that do exhibit duality do not display any statistics dependence of exponents at all.)

We should mention here also the work of Lütken and Ross³⁰ who were concerned with a description of QH systems on the basis of actions in which the scale dependent resistivities appear as parameters (along the lines of localization theory⁷) and postulated that a complexified duality or modular invariance operates on them. Their *ansatz* is equivalent to that of KLZ for our purposes.

Particle-hole symmetry. Within the framework of the fermionic Chern-Simons theory, there are analogous issues to the ones discussed above. However it is possible to gain some insight by studying the problem of noninteracting electrons in a random potential which is already nontrivial and at the very least may be a solution (in the Hartree-Fock sense) of the *interacting* $\nu=1$ to insulator transition, as suggested by some theoretical work.³¹

In particular, one can check if microscopic asymmetries of the random potential are irrelevant at the fixed point for the problem. In some measure this was done by Huo, Hetzel, and Bhatt³² in their numerical studies which found that critical conductivities were insensitive to departures from microscopic particle-hole symmetry. The contrast between their calculations of the density of states, a microscopic quantity,

and the long wavelength Hall conductivity (see Figs. 2 and 3 in their paper) also supports the conclusion that departures from particle-hole symmetry are indeed irrelevant. A second piece of evidence to this effect is the structure of the network model of Chalker and Coddington³³ which is believed to be in the same universality class. The network model clearly exhibits a symmetry between the QH and insulating phases; in the absence of boundary effects inessential to the bulk physics, they are chirally reversed translates of each other.

Constancy of ρ_H . The possibility of a quantized Hall resistivity into the insulator was implicit in the work of Jain and Kivelson²⁶ and KLZ who first suggested the idea of a Hall insulator where ρ_H would be finite. This was later studied in detail by Dykhne and Ruzin¹⁶ and Ruzin and Feng.¹⁷ In this work it is phrased in terms of a “semicircle law” relating $\sigma_L(\nu)$ and $\sigma_H(\nu)$. In the special case of a transition from a $1/k$ QH state to an insulator, this law states

$$\sigma_L^2 + \sigma_H^2 = \sigma_H/k \quad (28)$$

and is equivalent to the constancy of ρ_H , while in other cases it reflects the constancy of ρ_H for the “upper” fluid defined in the next section. Equation (28) was proved both for a classical two fluid model and within the network model, assuming linear response. Recently, it has been shown in a classicized (i.e., noninterfering) version of the network model that the constancy can persist beyond the linear response regime.³⁴ As argued in previous chapters, in terms of composite bosons this semicircle law corresponds to a vanishing Hall resistance. It is enlightening to point out, that a proof of $\rho_H^b=0$ is in fact included in Ref. 17, though in a different language. In their formulation, they introduced *local* current densities, \mathbf{j}_1 and \mathbf{j}_2 , for the two phases in the transition region and showed that their average values are perpendicular to each other—a property that is proved necessary and sufficient for the semi-circle law to be obeyed. Translating this to the bosonic Chern-Simons representation, liquids 1 and 2 correspond to mobile bosons and vortices, hence \mathbf{j}_1 and \mathbf{j}_2 are the charge and vortex current densities. The latter has the significance of an electric field in the perpendicular direction; Ruzin and Feng’s statement therefore implies that the current and voltage in the bosonic description are *parallel*—i.e., the Hall coefficient vanishes.

Finally, we should note that for the similar problem of the field tuned superconductor/insulator transition it has been argued by Fisher,³⁵ that an asymptotic particle-hole symmetry at the critical point might lead to a vanishing Hall coefficient as suggested by some data.³⁶ This does suggest the possibility that, in the QH system, both duality and the vanishing Hall response might ultimately be consequences of the same underlying principle.

V. DUALITY NEAR OTHER QUANTUM HALL PHASE TRANSITIONS

The hierarchical principle was invoked in Ref. 2 to argue that *all* QH transitions are, in a precise sense, transitions from principal QH states to insulators. Qualitatively, they all consist of a “lower” (parent) fluid that is inert across the transition and an “upper” (quasiparticle) fluid that undergoes a transition to an insulator. For example, the $\nu=2 \rightarrow 1$ transition is the $\nu=1 \rightarrow 0$ of the spin down lowest Landau

level while the spin up Landau level remains inert.

It follows that the symmetries of the transport observed in Ref. 1 should be present at all transitions if one can identify the transport coefficients of the upper fluid. This is straightforward for the conductivities. As the fluids conduct in parallel,

$$\begin{aligned}\sigma_L(E, \nu) &= \sigma_L^u(E, \nu), \\ \sigma_H(E, \nu) &= \sigma_H^u(E, \nu) + \sigma_H^l,\end{aligned}\quad (29)$$

where the lower fluid contributes only a fixed, quantized Hall conductivity independent of filling fraction and electric field.³⁷ As typical Hall bar measurements yield resistivities and, more importantly, the symmetry itself would be manifest in the resistivities of the upper fluid we need a prescription to go between them. In linear response this is simple: we convert the measured ρ_L and ρ_H to σ_L and σ_H by matrix inversion, obtain σ_L^u and σ_H^u from Eq. (29), and invert *them* to get ρ_L^u and ρ_H^u .

Beyond linear response one needs explicit expressions for the current carried by the upper fluid, which is *not* parallel to the net current, and for the electric fields resolved along and perpendicular to it. Consider a Hall bar geometry where a current density j flows along the bar. Voltage measurements yield the longitudinal and Hall electric fields E_L, E_H and hence the (in general) nonlinear resistivities $\rho_L(j), \rho_H(j)$. Let us denote the Hall angle between the current and the total electric field of magnitude $E = \sqrt{E_L^2 + E_H^2}$ by θ , so that $E_L = E \cos \theta, E_H = E \sin \theta$. Denoting the current density and Hall angle for the upper fluid by j^u and θ^u , Eq. (29) can be recast as

$$\begin{aligned}j \cos \theta &= j^u \cos \theta^u, \\ -j \sin \theta &= -j^u \sin \theta^u + \sigma^l E.\end{aligned}\quad (30)$$

The two equations in Eq. (30) suffice to determine j^u and θ^u , and hence E_L^u and E_H^u , in terms of the applied j and measured ρ_L and ρ_H . After some algebra we obtain,

$$\begin{aligned}j^u &= j \left\{ \frac{\rho_L^2 + [\rho_H + \sigma_H^u(\rho_L^2 + \rho_H^2)]^2}{\rho_L^2 + \rho_H^2} \right\}^{1/2}, \\ E_L^u &= \frac{j^2}{j^u} \rho_L, \\ E_H^u &= \frac{j^2}{j^u} [\rho_H + \sigma_H^u(\rho_L^2 + \rho_H^2)]^2.\end{aligned}\quad (31)$$

Plots of E_L^u and E_H^u as functions of j^u should then be expected to resemble the I - V_L, I - V_H characteristics reported in Ref. 1.

VI. SUMMARY AND OPEN QUESTIONS

In this paper we have followed a single line of argument in interpreting the experimental results, i.e., we have assumed that the physics in the transition region is governed by a zero temperature quantum critical point. A great virtue of such an interpretation is that measured quantities become properties of a scaling limit where it is possible for symme-

tries, not manifest microscopically, to emerge because the operators that break them are irrelevant at the underlying fixed point. In other words, critical points can provide a robust rationale for long wavelength symmetries. Nevertheless, we should remark that this is not the only possibility. While the asymptotic low temperature region (if it can be accessed on realistic time scales) for large samples should be governed by critical physics one has to leave open the possibility that the accessible temperature range might involve more complicated finite temperature effects and lead to some of the same physics for more classical reasons as in Refs. 16,34. For the samples studied to date that exhibit the reflection symmetry, the temperature range that shows evidence for scaling is too small for us to rule out such a possibility.

With this caveat, we have argued that the reflection symmetry can be naturally interpreted as duality/particle-hole symmetry combined with a vanishing/constant Hall response in the composite boson/fermion descriptions of the QH to insulator transitions. We expect that this is a feature of all continuous QH transitions and have indicated how to search for it at other transitions. We have also reviewed the theoretical evidence in support of these inferences and find that though not dispositive, it is certainly encouraging; at least for noninteracting electrons, it is easy to visualize calculations that can test them further.²⁷ Overall, we find that the experimental data through the interpretation of duality offer the strongest evidence yet that the general framework of KLZ of a universal bosonic transition underlying the QH transitions is correct.

The fit between our analysis and the data, though compelling, is not perfect. In particular, the derivation of the relationship between ν_d and ν should be strictly valid only for a system that exhibits duality down to the microscopic length scale, e.g., for noninteracting electrons this is the case for microscopically particle-hole symmetric disorder in the lowest Landau level. For systems where duality is recovered only at long wavelengths, the relationship should be more complicated. Indeed, in the strict scaling limit, T and $\nu - \nu_c \rightarrow 0$ with $x = (\nu - \nu_c)/T^{1/\nu_z}$ fixed, it matters very little what we pick. Nevertheless, this does *not* mean that duality is without consequence. For example, it would still imply that, as a function of the scaling variable x , the current and voltage trade places at dual values and would have the consequence that $\rho_L^c = 1$. We feel that the correct perspective on our analysis is that we have approximately identified a nonlinear scaling field ($\nu' - \nu'_c$ in the fermionic description) that allows the symmetry to be identified over a wider range of ν at accessible values of T . It would however be very useful to get some quantitative understanding of why $\nu' - \nu'_c$ continues to be such a good scaling field even when ν'_c itself is shifted from its symmetric value of 1/2 by as much as 20%. This problem becomes more serious if the transition out of the QH state is studied as a function of disorder at fixed filling. While the formulation in terms of the scaling variable remains valid it is not obvious how one might interpret any data, away from the scaling limit, that might become available for this transition. A second problem, which is difficult to pin down experimentally given the difficulty of accurately measuring ρ_H deep into the insulator, is that the reflection symmetry for the longitudinal response appears to hold over a larger range of fillings than those over which the Hall re-

sponse is constant. We do not have a good understanding of this difference.

A different issue is the nature of the nonlinear response. We have suggested that the nonlinear response is a consequence of intrinsic, critical physics. It is not difficult to imagine getting the same result from a heating scenario in which the electron gas equilibrates at a different temperature from the lattice. What one needs for this purpose is the reflection symmetry of the linear response plus an effective electronic temperature that depends upon the dissipation in the bulk alone, i.e., on the product jE_L which is the same between dual points on the longitudinal characteristics. The reason we are suspicious of this mechanism is that it ignores the dissipation at the contacts, where the Hall voltage is dropped and which is very asymmetric between dual points. This is clearly an important issue and we expect to make progress on it by further analysis of the data in the near future. We should note that even if a heating scenario is correct, the symmetries of the linear response would still require an explanation.

Finally, we would like to draw attention to the possibility of analogous symmetries near other interesting transitions. The most closely related one is the field-tuned transition in two-dimensional disordered superconducting films.¹⁹ Here

the theoretical expectation¹⁵ is that it should *not* be characterized by duality as the bosons and vortices interact with different potentials. Unfortunately, the data here is not completely consistent, particularly on the question of universal values for the critical resistivities; further studies with an emphasis on the current-voltage characteristics might therefore be quite useful. There is also the observation of a reflection symmetry at the puzzling zero-field transition observed by Kravchenko *et al.*³⁸ which does not have any natural interpretation in our framework. We expect these cross comparisons to be very instructive in evaluating the correctness of the analysis outlined in this paper.

ACKNOWLEDGMENTS

We thank A. Auerbach, Y. Avron, M. P. A. Fisher, E. Fradkin, D. E. Freed, S. M. Girvin, S. A. Kivelson, M. Stone, and D. C. Tsui for instructive discussions. E.S. is grateful for the hospitality and support of the Physics Department and ITP at the Technion, Israel, where part of this work was carried out. This work has been supported by the Beckman Foundation, the Aspen Center for Physics, the Israeli Academy of Sciences (ES), the A. P. Sloan Foundation, NSF Grant No. DMR-9632690 (S.L.S.) and the NSF (D.S.).

- ¹D. Shahar, D. C. Tsui, M. Shayegan, E. Shimshoni, and S. L. Sondhi, *Science* **274**, 589 (1996).
- ²S. Kivelson, D. H. Lee, and S. C. Zhang, *Phys. Rev. B* **46**, 2223 (1992).
- ³S. M. Girvin and A. H. MacDonald, *Phys. Rev. Lett.* **58**, 1252 (1987); S.-C. Zhang, T. H. Hansson, and S. Kivelson, *ibid.* **62**, 82 (1989); N. Read, *ibid.* **62**, 86 (1989); A. Lopez and E. Fradkin, *Phys. Rev. B* **44**, 5246 (1991); R. Rajaraman and S. L. Sondhi, *Int. J. Mod. Phys. B* **10**, 793 (1996).
- ⁴J. K. Jain, *Phys. Rev. Lett.* **63**, 199 (1989).
- ⁵Different formulations of the hierarchy are due to F. D. M. Haldane, *Phys. Rev. Lett.* **51**, 605 (1983); B. I. Halperin, *ibid.* **52**, 1583 (1984); **52**, 2390(E) (1984); and J. K. Jain Ref. 4. Their long wavelength equivalence was established by N. Read, *Phys. Rev. Lett.* **65**, 1502 (1990).
- ⁶For a summary of departures from KLZ's original phase diagram see S. H. Song, D. Shahar, D. C. Tsui, Y. H. Xie, and Don Monroe, *Phys. Rev. Lett.* **78**, 2200 (1997).
- ⁷See A. M. M. Pruisken, in *The Quantum Hall Effect*, edited by R. E. Prange and S. M. Girvin (Springer-Verlag, New York, 1986).
- ⁸H. P. Wei, D. C. Tsui, M. A. Paalanen, and A. M. M. Pruisken, *Phys. Rev. Lett.* **61**, 1294 (1988).
- ⁹Transitions between fractional states were first studied by L. W. Engel, H. P. Wei, D. C. Tsui, and M. Shayegan, *Surf. Sci.* **229**, 13 (1990), and between QH states and the insulator by L. W. Wong, H. W. Jiang, N. Trivedi, and E. Palm, *Phys. Rev. B* **51**, 18 033 (1995).
- ¹⁰The experimental situation is summarized in S. L. Sondhi, S. M. Girvin, J. P. Carini, and D. Shahar, *Rev. Mod. Phys.* **69**, 315 (1997).
- ¹¹See B. Huckestein, *Rev. Mod. Phys.* **67**, 357 (1995), and references therein.
- ¹²X. G. Wen and Y. S. Wu, *Phys. Rev. Lett.* **70**, 1501 (1993); W.

- Chen, M. P. A. Fisher, and Y. S. Wu, *Phys. Rev. B* **48**, 13 749 (1993).
- ¹³L. P. Pryadko and S. C. Zhang, *Phys. Rev. Lett.* **72**, 3282 (1994); *Phys. Rev. B* **54**, 4953 (1996); E. Fradkin and S. Kivelson, *Nucl. Phys. B* **474**, 543 (1996).
- ¹⁴D. Shahar, D. C. Tsui, M. Shayegan, J. E. Cunningham, E. Shimshoni, and S. L. Sondhi, *Solid State Commun.* (to be published).
- ¹⁵M. P. A. Fisher, *Phys. Rev. Lett.* **65**, 923 (1990).
- ¹⁶A. M. Dykhne and I. M. Ruzin, *Phys. Rev. B* **50**, 2369 (1994).
- ¹⁷I. M. Ruzin and S. Feng, *Phys. Rev. Lett.* **74**, 154 (1995).
- ¹⁸S. C. Zhang, T. H. Hansson, and S. A. Kivelson, *Phys. Rev. Lett.* **62**, 82 (1989).
- ¹⁹Field-tuned superconductor to insulator transitions have been observed in (a) disordered films by A. F. Hebard and M. A. Paalanen, *Phys. Rev. Lett.* **65**, 927 (1990) and A. Yazdani and A. Kapitulnik, *ibid.* **74**, 3037 (1995); (b) in Josephson junction arrays by H. S. J. van der Zant, F. C. Fritschy, W. J. Elion, L. J. Geerligs, and J. E. Mooij, *ibid.* **69**, 2971 (1992); and (c) in the cuprates by G. T. Seidler, T. F. Rosenbaum, and B. W. Veal, *Phys. Rev. B* **45**, 10 162 (1992). Reviews of field and disorder tuned zero-field superconductor/insulator transitions in two dimensions are A. F. Hebard, in *Strongly Correlated Electronic Materials*, The Los Alamos Symposium 1993, edited by K. S. Bedell, Z. Wang, D. E. Meltzer, A. V. Balatsky, and E. Abrahams (Addison-Wesley, Reading, MA, 1994), p. 251; G. T. Zimanyi, *ibid.* p. 285; Y. Liu and A. M. Goldman, *Mod. Phys. Lett. B* **8**, 277 (1994).
- ²⁰Duality and self-duality are often used in the literature in somewhat different ways. The dual representation of a classical or quantum statistical mechanics is one that interchanges the roles of the ordered and disordered phases but does not, in general, involve the same degrees of freedom, e.g., the dual of the three-dimensional Ising model is an Ising gauge theory. When the

- degrees of freedom are the same and especially when the Hamiltonians are of the same form, it is sometimes said that the system exhibits self-duality, e.g., the two-dimensional Ising model. We refer to such a case as duality, reserving the term self-duality for the critical point where the dual formulation is identical to the original one, e.g., the two-dimensional Ising model at its critical point. For an introduction to duality in statistical mechanics see, R. Savit, *Rev. Mod. Phys.* **52**, 453 (1980).
- ²¹The principle of resistivity addition in problems with gauge fields enforcing constraints was recognized by L. B. Ioffe and A. I. Larkin, *Phys. Rev. B* **39**, 8988 (1989).
- ²²Here F is a functional of a dynamic field, hence the suggestion of a free energy is not appropriate. However, as it plays the role of a dynamical analog of the equilibrium energy functional, we hereon use this term and drop the quotation marks.
- ²³E. Shimshoni and S. L. Sondhi (unpublished).
- ²⁴M. P. A. Fisher and D. H. Lee, *Phys. Rev. B* **39**, 2756 (1989); D. H. Lee and M. P. A. Fisher, *Int. J. Mod. Phys. B* **5**, 2675 (1991).
- ²⁵J. K. Jain, S. A. Kivelson, and N. Trivedi, *Phys. Rev. Lett.* **64**, 1297 (1990); **64**, 1993(E) (1990).
- ²⁶J. K. Jain and S. A. Kivelson, *Phys. Rev. B* **37**, 4276 (1988).
- ²⁷A. Lopez, J. T. Chalker, and S. L. Sondhi (unpublished).
- ²⁸If the ideal regions have nonlinear dispersions then the Landauer conductances are not properties of the disordered regions alone and the scheme is no longer sensible for our purposes.
- ²⁹E. Chow, H. P. Wei, S. M. Girvin, and M. Shayegan, *Phys. Rev. Lett.* **77**, 1143 (1996).
- ³⁰C. A. Lutken and G. G. Ross, *Phys. Rev. B* **48**, 2500 (1993).
- ³¹D. H. Lee and Z. Q. Wang, *Phys. Rev. Lett.* **76**, 4014 (1996).
- ³²Y. Huo, R. E. Hetzel, and R. N. Bhatt, *Phys. Rev. Lett.* **70**, 481 (1993).
- ³³J. T. Chalker and Coddington, *J. Phys. C* **21**, 2665 (1988).
- ³⁴E. Shimshoni and A. Auerbach (unpublished).
- ³⁵M. P. A. Fisher, *Physica A* **177**, 553 (1991).
- ³⁶M. A. Paalanen, A. F. Hebard, and R. R. Ruel, *Phys. Rev. Lett.* **69**, 1604 (1992).
- ³⁷There are certainly deviations from the ideal response even for an ideal QH state, but they are far smaller than the conductivities of the upper fluid and can be neglected in comparison.
- ³⁸S. V. Kravchenko, D. Simonian, M. P. Sarachik, Whitney Mason, and J. E. Furneaux (unpublished); D. Simonian, S. V. Kravchenko, and M. P. Sarachik (unpublished).