1-1-1995

Nonlinear Steady-State Mesoscopic Transport, I. Formalism

M. D. Johnson
University of Central Florida

O. Heinonen
University of Central Florida

Nonlinear steady-state mesoscopic transport: Formalism

M.D. Johnson and O. Heinonen
Department of Physics, University of Central Florida, Orlando, Florida 32816-2385
(Received 1 February 1995)

We present an approach to steady-state mesoscopic transport based on the maximum entropy principle formulation of nonequilibrium statistical mechanics. Not restricted to linear response, this approach is valid in the nonlinear regime of high current, and yields the quantization observed in the integer quantum Hall effect at large currents. A key ingredient of this approach, and its success in explaining high-precision Hall measurements, is that the occupancy of single-electron states depends on their current as well as their energy. This suggests that the reservoir picture commonly used in mesoscopic transport is unsatisfactory outside of linear response.

INTRODUCTION

We have recently developed an approach to steady-state mesoscopic transport not restricted to the linear response regime. By nonlinear here we mean that the driving force (voltage or current) is large—too large for linear response calculations, but well within the range of typical experimental values. Our approach is applicable to quasi-one-dimensional systems and to two-dimensional systems in strong magnetic fields, such as those exhibiting the integer quantum Hall effect. This work was reported in a brief form elsewhere.[1] Our study was motivated by the important but often neglected fact that the integer quantum Hall effect[2] (IQHE) is exhibited even when the currents and voltages are very large.[3] Because the magnetic field strongly suppresses inelastic scattering, systems exhibiting the IQHE can be viewed as mesoscopic even though they are relatively large,[4] and the IQHE itself can be viewed as a near-ideal manifestation of mesoscopic transport. The IQHE at low currents is well understood in terms of the Landauer-Büttiker (LB) approach[5,6] to mesoscopic transport. This approach, however, appears to be fundamentally a linear response theory.[7,8] (See, however, Refs. 9 and 10.) When applied nonetheless at high currents, it fails to yield the observed quantization. The quantization at high currents appears to us to be an extraordinary phenomenon. It is easy to imagine many things that can end quantization (dissipation, backscattering, etc.), but it is not obvious how to restore quantization. It is perhaps possible that quantization at high current might result from conventional approaches to nonequilibrium transport (such as the Keldysh or quantum Boltzmann approaches[11]); but these are difficult even near linear response and their behaviors at large currents simply unknown. The high-current quantization is so extraordinary that it seemed likely to us that a successful theory of large-current mesoscopic transport would have to take its highly nonequilibrium nature into account from the very beginning.

We found an apparently fruitful direction in the maximum entropy approach (MEA) to nonequilibrium statistical mechanics,[12] in which the density matrix is found by maximizing the information entropy of the system, subject to constraints which fix the expectation values of observables. This approach should in principle be quite generally applicable to equilibrium and nonequilibrium statistical mechanics, but in fact there have been very few examples of the latter. For reasons which we shall explain below, mesoscopic systems (including those exhibiting the IQHE) are ideally suited for the MEA, we will present calculations that are exact for mesoscopic systems consisting of noninteracting electrons.[13] This gives as a first benefit an exact Hall quantization at zero temperature, for ideal systems, for almost arbitrarily large currents. A consequence of this work is that it suggests that the picture of current and voltage probes as *reservoirs*—which underlies nearly all approaches to transport in mesoscopic systems, and which has proven extremely successful at low currents—may not be satisfactory at high currents. Finally, our approach also establishes a connection between the gauge argument originally proposed by Laughlin[14] for the IQHE in closed systems and the IQHE in open systems.

In the present and a companion paper we present our theory in detail together with discussions and applications. The present paper, part I, develops the formal theory in the ideal case, and the companion, part II, contains various applications. In the present paper, Sec. I contains a more detailed introduction and motivation. Section II contains our maximum entropy approach together with a detailed example, and Sec. III contains a discussion of the maximum entropy approach, the resulting electronic distributions, and the relationship between this work and more conventional approaches.

I. LANDAUER-BÜTTIKER FORMALISM

A mesoscopic system typically consists of a "device" (such as a Hall bar or quantum wire) and a number of current and voltage probes. The device itself is truly mesoscopic, by which we mean that the length of the device is smaller than the phase-breaking length. The connection between the device and the external world is

provided by the probes, which are attached to the device at terminals. (In reality, the terminals consist of, for example, the n^+-doped regions which connect to the inversion layer in a silicon-metal-insulator field-effect transistor, or the "fingers" of indium diffused through the different layers in a heterostructure. Thus, in a real device carriers pass into the device through terminals which are no larger than the device itself.) The carriers in general will suffer both elastic and inelastic scattering, and so dissipate energy, in the terminals. We presume that within the mesoscopic device itself scattering is entirely elastic and dissipationless.

It is not obvious how to handle the complicated system of device plus probes. The greatest advance in understanding mesoscopic transport came from an approach originally due to Landauer,[5,10,15,6] in which the conduction of the entire sample is treated as a scattering problem. There are two central concepts in this model. The first is the "reservoir"— the probes are treated as macroscopically large (essentially infinite) reservoirs which inject carriers into the device through ideal leads. It is assumed that each reservoir is in local equilibrium described by a local chemical potential μ_m, and that the occupancy of electronic states entering the device from a reservoir is determined by the local chemical potential of that reservoir. We will use the term "reservoir" only in this restricted sense. More generally we will refer to a source of carriers as a "terminal." The second key concept is that the motion of carriers through the device itself is treated as an elastic scattering problem. Carriers entering from a reservoir into the device are scattered either back into the original reservoir or outward into the other reservoirs. The scattered electrons then equilibrate deep within the reservoir with the electrons in the reservoir. Scattering in the terminals randomizes the energy and phase of the carriers, which eliminates any quantum interference.[16]

To formulate the scattering problem one needs asymptotic regions in the leads in which states carry electrons either away from or towards the device[16] (and in which evanescent modes have decayed away). Such leads are quasi-one-dimensional, and states in them can be labeled by subband index n and wave number k. (In the presence of a magnetic field, n is a Landau level index.) Within the system consisting of the device plus the asymptotic regions scattering is elastic. This gives the conventional treatment of mesoscopic devices: conduction occurs as elastic scattering of carriers injected into the asymptotic regions from reservoirs.[16] In this paper we will adhere entirely to the scattering viewpoint. We will argue, however, that outside the linear regime one cannot treat the terminals (which inject carriers) as reservoirs in the specific sense defined above.

To be definite, we will use the following notation and conventions. For simplicity, we assume that the electrons are noninteracting and spinless. (Both restrictions can be removed—spin just adds another label, and the formulation we give can be extended to the case of interacting many-body states.) In an M-terminal device we denote the probes by $m = s, d, 2, \ldots, M-2$. Here s denotes the source of current and d the drain, since in typical experiments current flows in one lead and out one other. In general we denote a complete orthogonal set of single-particle eigenstates by $|\psi_\alpha\rangle$. These have energies ϵ_α and carry net currents $i_{m,\alpha}$ through terminal m. The scattering picture mentioned above can be made precise by supposing that the leads can be treated as semi-infinite and straight.[16] In this case, a particularly useful set of eigenstates for multiterminal ($M > 2$) systems are the scattering states[8,17] $|\psi^+_{mnk}\rangle$ (i.e., $\alpha = mnk$). The state $|\psi^+_{mnk}\rangle$ is incoming into the device from terminal m; n and k denote the asymptotic wave number and subband index of the incoming wave. The state has energy ϵ_{mnk} and carries current i^0_{mnk} into the device at terminal m. (Landauer objects to this lead geometry as incompatible with the reservoir concept.[10] But our discussion of the LB approach needs only a subband index and wave number, and makes no other use of this geometry. We have specified this geometry here because the scattering language it makes precise is useful for the subsequent sections.) In general, after scattering within the device, the state carries outward current through each terminal. We denote by $i_{m',mnk}$ the *net* current of this state into the device at terminal m', with the convention that current flowing *into* the device is positive. The state's net current at m' is related to its incoming current at m by $i_{m',mnk} = i^0_{mnk}(\delta_{m'm} - \sum_{n'k'} T_{m'n'k',mnk})$, where $T_{m'n'k',mnk}$ is the transition probability obtained from the scattering matrix in the $|\psi^+_{mnk}\rangle$ representation (with a proper normalization[17,18]). The elastic scattering within terminals which can exist in a real system can be included as contributions to the transmission probabilities. We will occasionally use second quantization in which the electron field operator $\hat{\psi}(\mathbf{r})$ has an expansion in the states $\psi^+_{mnk}(\mathbf{r})$ given by

$$\hat{\psi}(\mathbf{r}) = \sum_{mnk} \psi^+_{mnk}(\mathbf{r})\hat{c}_{mnk}, \tag{1}$$

where $\hat{c}_{mnk}$ destroys a state[19] $\psi^+_{mnk}(\mathbf{r})$.

The reservoir and scattering concepts underlie the Landauer-Büttiker (LB) theory of mesoscopic transport.[5,6,16] A key point in this theory is that voltage as well as current contacts are treated identically and described as reservoirs.[6] The reservoirs enter this theory in several important ways: they determine the electronic distribution, they randomize the phase of occupied states (which eliminates interference), and they provide a prescription for determining voltage differences. Combined, these permit one to calculate current-voltage (I-V) curves. First consider the distribution of electrons in the device. Suppose the mth terminal is a reservoir described by a local chemical potential μ_m. Then the states in the attached lead carrying current toward the device are occupied according to the Fermi-like distribution $f^{\rm LB}_{mnk} = 1/[e^{\beta(\epsilon_{mnk}-\mu_m)} + 1]$. (Here $\beta = 1/k_BT$, where T is the temperature and k_B is Boltzmann's constant.) The net current flowing into the system at, say, the source is $I = \sum_{mnk} i_{s,mnk} f^{\rm LB}_{mnk}$. Suppose that in a multiprobe device the current flowing in at the source and out at the drain is I, with zero net current at other terminals. In the LB approach this is enough information to

determine the local chemical potentials μ_m (within an additive constant). Within the reservoir picture, it is then obvious that the measured voltage difference between two terminals (1 and 2, say) must be $V = (\mu_1 - \mu_2)/e$, where e is the charge of an electron.

Consider the particular case of a two-terminal device at zero temperature, with some current I flowing from source to drain. The current is in this picture caused by a voltage difference $V = (\mu_s - \mu_d)/e = IR$ between source and drain. Following the standard LB approach, at low voltages this gives a resistance $R = h/(je^2\widetilde{T})$, where j is the number of occupied subbands and $\widetilde{T}$ is the total transmission probability at μ_s. This resistance is quantized in the absence of backscattering ($\widetilde{T} = 1$). In real quantum Hall systems, there is a macroscopic separation between left- and right-moving states, so that in fact backscattering is highly suppressed, and indeed $\widetilde{T}$ is very nearly unity.[20] Here we have briefly given the two-terminal version of this explanation, but following Büttiker[6] it can be generalized to the multiterminal case (see below), in which case the corresponding voltage is the transverse or Hall voltage. (We have here neglected dissipation that occurs due to contact resistance at the source and drain. In a real two-terminal device this dissipation prevents perfect quantization of resistance. Resistances can be quantized only in a multiterminal device when measured between two terminals through which no net current flows.) Hence the LB theory gives a very satisfying microscopic explanation of how the conductance can be so accurately quantized in the IQHE (Ref. 16)—at least in the low-current regime, as we will explain below.

We have summarized the LB theory here in a way appropriate for linear response—the difference $\mu_s - \mu_d$ is presumed to be small. This is assumed in nearly all uses of the LB approach.[6,16,21] For example, we treated the transmission $\widetilde{T}$ as a constant, which requires in part that $\mu_s - \mu_d$ be small. It is possible to generalize this to the case where the transmission depends significantly on energy near the Fermi level,[22] as would be needed if $\mu_s - \mu_d$ is large. In this paper, however, we will be concerned only with ideal systems (perfect transmission), so this type of generalization is not relevant here.

The LB theory of macroscopic transport has been used to interpret a wide variety of experiments (see, for example, a partial listing in Ref. 21). In fact, the fundamental model of reservoirs at (local) equilibrium with local chemical potentials μ_m has been used in essentially all mesoscopic calculations, including microscopic methods such as nonequilibrium Green's functions. There is no reason to doubt the fundamental soundness of treating the terminals as reservoirs in the low-current regime.

Despite the many successes of the LB approach, there are important experiments which it seems unable to explain. Chief among these are the actual high-precision IQHE experiments. In our above application of the LB approach to the IQHE we did not mention an important point: the LB theory only predicts quantization in the IQHE when the current is small. When the current is large (as large as those typical in high-precision experiments), the same argument predicts a failure of quantization. There is at least no straightforward way to extend the LB approach to give quantization at large currents. The culprit appears to be the very reservoir concept itself.

Let us examine this in detail, first in an ideal two-terminal example. Denote the minimum energy of the jth subband by E_j. Suppose to begin with that the local chemical potentials of both the source and drain exceed E_j but are less than E_{j+1}. [See Fig. 1(a).] In an ideal system at zero temperature, the net current in each subband is, according to the LB approach, $e(\mu_s - \mu_d)/h$. (This simple form occurs because of a cancellation between the density of states and the current carried in each single-particle state.[23]) With j occupied bands, this gives a current $I = je(\mu_s - \mu_d)/h$. The two-terminal voltage in the reservoir picture is $V = (\mu_s - \mu_d)/e$, and so the two-terminal resistance is $R = V/I = h/(je^2)$.

Now let us suppose that the local chemical potential of the source (but not of the drain) is increased above E_{j+1}, so that the source injects electrons into the $(j+1)$st subband, while the drain does not. This is shown in Fig. 1(b). In an ideal system at zero temperature, the net current contributed by each of the lower j subbands is the same as above, but the net current in the $(j+1)$st

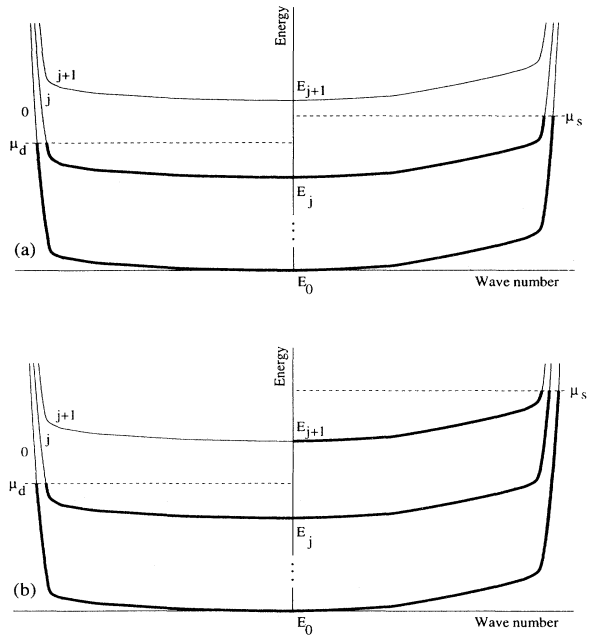

FIG. 1. Schematic energy spectrum of a one-dimensional mesoscopic device, or a laterally confined two-terminal IQHE system, and zero-temperature occupancies of the single-particle states in the Landauer-Büttiker picture. In this picture the occupancies of the states entering the device from the source and drain are described by local chemical potentials μ_s and μ_d, respectively. Occupied states are marked by a heavy line. The asymmetric band bending shown could arise from electrostatics, such as a combination of Hall field and confinement potential in a Hall device. In (a) $\mu_{s,d}$ exceed the minimum of the lowest j subbands (or Landau levels), but lie below the minimum of the $(j+1)$st subband. In (b) μ_s (but not μ_d) lies above the minimum of the $(j+1)$st subband.

subband is $e(\mu_s - E_{j+1})/h$. Thus the total current is $I = je(\mu_s - \mu_d)/h + e(\mu_s - E_j)/h$. The voltage difference must still be $V = (\mu_s - \mu_d)/e$ in the reservoir picture, so in this case the LB approach gives a two-terminal resistance $R = V/I$ which lies between $h/(je^2)$ and $h/[(j+1)e^2]$. Notice that this happens whenever one or more of the subband minima lie between μ_d and μ_s, which must occur whenever eV exceeds the subband spacing. That is, even in the case of an ideal system, with no backscattering, the LB prediction is that the two-terminal resistance should not be quantized when eV is large. And yet in the high-precision IQHE experiments quantization is found to be extremely accurate even when eV is 10 or 100 times the subband spacing.[3] (The same argument has been invoked to explain the large-voltage failure of resistance quantization in quantum point contact experiments within the LB formalism.[24])

The above argument also works in an ideal multiterminal IQHE system. Suppose there are four terminals (taken to be identical for simplicity), with two current terminals s and d, and two transverse voltage terminals 1 and 2 (see Fig. 2). Then to get zero net current in terminal 1, it is necessary that $\mu_1 = \mu_s$; similarly, $\mu_2 = \mu_d$. A given μ_s and μ_d give the same total source-to-drain current as above, and the transverse (Hall) voltage is $(\mu_1 - \mu_2)/e$, also as above. Hence the ideal four-terminal LB results are identical to the ideal two-terminal measurements: the resistances are quantized at low but not at high currents.

The discussion thus far has treated the electrons as noninteracting. With a large current the LB distribution would move many electrons from one edge to the other. Obviously this would have a big electrostatic effect which we have so far neglected but which might restore quantization. In fact, including electrostatics does not help. At the lowest (Hartree) level of approximation, electron-electron interactions can cause the subbands to deviate, perhaps significantly, from the wave number dependence which would arise in the noninteracting case.[23] This consequence of electrostatics is pictured schematically in Fig. 1, which shows the bending of energy levels (not too near a contact) in an IQHE sample due to the combined presence of a Hall field and an edge confinement. Of itself, this bending of bands has no effect on the argument above. If states in one level are occupied out to μ_s and μ_d, then this level contributes $e(\mu_s - \mu_d)/h$ to the current, regardless of the band's shape (because of the cancellation between density of states and current per state). So as long as all subbands shift in energy more or less together, as is the case with Hartree interactions, the picture is unchanged: a large voltage causes partial occupancy of higher subbands and hence failure of quantization. (We will in part II present detailed numerical calculations of the electrostatic fields in an ideal Hall bar with a large current.)

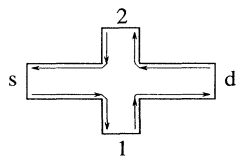

FIG. 2. Schematic representation of a typical four-terminal IQHE device. Current runs along the lower edge from s to terminal 1 and then to d. Similar currents run along the upper edge. If the device is ideal, and all terminals are identical, then in the Landauer-Büttiker picture $\mu_1 = \mu_s$ and $\mu_2 = \mu_d$ to get zero net current in terminals 1 and 2. In the maximum entropy approach under the same circumstances $\xi_1 = \xi_s$ and $\xi_2 = \xi_d$.

Van Son and Klapwijk recently attempted to examine more closely the consequences of electrostatics for the LB approach to the IQHE.[9] Their starting point is that electron states (except very close to the current probes) can be described by bulk Landau levels with a shape given by a self-consistent electrostatic potential. They argue that the source injects electrons only in a range of energy $\mu_s - \Delta < \epsilon < \mu_s$, but that near the source these "relax" to fill (bulk) subbands up to some energy μ_s' (with $\mu_s' < \mu_s$). For example, when $\sigma_H = e^2/h$, after relaxation the lowest subband is filled from μ_d (for left movers) to $\mu_s' < \mu_s$ (for right movers). Note that, if μ_s' lies above the minimum of the next subband, this "relaxation" results in some occupied states (e.g., in the lowest subband) lying at higher energies than empty states in the next subband. There are two problems with this analysis. First, in a multiterminal device this distribution would be changed at the first voltage probe downstream in some unknown way which one should expect would prevent quantization. Second, Van Son and Klapwijk simply assume that the relaxation process is so efficient that at low temperatures states in the lowest subband end up filled continuously from μ_d to μ_s' (even though the current source and drain do not directly feed all of them), while the upper subbands all are empty (for $\sigma_H = e^2/h$). It is difficult to believe that no carriers would end up in any of the higher subbands for large voltages, where $\mu_s - \mu_d$ is 10 or 100 times $\hbar\omega_c$. Moreover, this perfect relaxation would have to suddenly (and inexplicably) end as soon as the drain also begins inserting carriers directly into the next Landau level, if this is also to explain the IQHE at integers greater than 1.

Let us turn back to the usual reservoir picture, and examine whether including the interactions more accurately—by, say, including exchange—can restore the high-current quantization. The exchange interaction effectively lowers the energy of the occupied single-particle states relative to unoccupied states, and in principle this could remove the partial occupancy of higher subbands (the cause of the failure to quantize). But the IQHE quantization at large voltages[3] would require the exchange and correlation energies to exceed $10\hbar\omega_c$ or even $100\hbar\omega_c$. In fact these energies are much smaller (of order $10^{-3}\hbar\omega_c$), and it seems implausible that interactions could restore quantization.

Landauer's insight was that it should be possible to ignore the (perhaps very complicated) details of the terminals, which are described entirely by the transmission probabilities, and concentrate on general principles (see, for example, the discussion in Ref. 10). The existence of high-current quantization in the IQHE and other meso-

scopic devices argues, we believe, that this insight is fundamentally correct. However, based on the experimental evidence of the high-precision measurements of the quantum Hall resistance, we believe that the model of terminals as macroscopic reservoirs in local equilibrium is unsatisfactory outside the regime of linear response. Here we will present an approach to mesoscopic transport which can be viewed as being within the spirit of Landauer's idea, in that we assume it is possible to neglect the details of the terminals. Our approach leads, however, to a different occupancy of electronic states (and hence requires modification of the reservoir concept), and appears capable of describing transport in mesoscopic devices outside of the low-current realm.

II. MAXIMUM ENTROPY APPROACH TO STEADY-STATE MESOSCOPIC TRANSPORT

We start our approach to steady-state mesoscopic transport with the following fundamental assumption. We assume that the steady state of a mesoscopic system can be described in terms of an ensemble of electron (or more generally, carrier) states in the mesoscopic device itself. This is analogous to standard assumptions in quantum statistical mechanics in which a system coupled to a heat bath or a particle reservoir can be described by an ensemble average over *closed* systems in which the coupling between the particles in the systems and the heat bath or reservoir is not explicitly included, but enters only implicitly through Lagrangian multipliers and constraints. Here, we consider the electron states in a system with either open or periodic boundary conditions, and do not explicitly include a coupling between the electrons and dissipative processes in the terminals. We will also ignore evanescent modes which decay exponentially from the terminals.

We will use the maximum entropy approach to nonequilibrium statistical mechanics to obtain the distributions over the ensemble of electrons. In this section we will develop our approach to steady-state mesoscopic transport and provide a detailed example which illustrates how it can explain the IQHE even at large currents and voltages.

A. Density matrix and current operator

The central ingredient of the maximum entropy approach (MEA) is the information entropy[12] S_I. If a complete set of eigenstates of a thermodynamic system is labeled by γ, then the information entropy is given by

$$S_I = -c \sum_{\gamma} P_\gamma \ln P_\gamma \ , \tag{2}$$

where P_γ is the probability that the system is in a given microstate $|\gamma\rangle$. Here c is an unspecified constant. (When the MEA is applied to equilibrium thermodynamics, it can be shown that the information and thermodynamic entropies are identical when c is chosen to be Boltzmann's constant k_B. We will see that this is also true in the case of steady-state mesoscopic transport.) As in any thermodynamic calculation, the first necessity is to define "the system." We assume that we can define the system to be the device (including the asymptotic leads), as described earlier. In general, γ then refers to one of a complete set of many-body electron states. In the case of noninteracting electrons each such microstate corresponds to a particular set of occupied single-particle scattering states. It is the fact that the device itself is mesoscopic which allows for a straightforward description of the microstates. In the presence of dissipation, this is not so easy. In fact, in spite of claims to its general applicability, nearly all applications of the MEA to dissipative nonequilibrium systems have been limited to expansions about equilibrium.[12]

The fundamental postulate of the MEA is that the probabilities P_γ are those which maximize the information entropy—subject to constraints which describe certain given or known observables. The method itself does not give a prescription for determining what are the constraints. These must be determined from physical considerations. (We will discuss this point in more detail later.) In the case of equilibrium thermodynamics, it is assumed that the internal energy U and electron number N can be taken as given, whether or not they are actually measured. In the case of steady-state transport one knows in addition (by measurement) the net current at each terminal. We therefore include this as an additional constraint and maximize the information entropy subject to the constraints $\langle \hat{H} \rangle = U$, $\langle \hat{N} \rangle = N$, and $\langle \hat{I}_m \rangle = I_m$. Here $\hat{H}$ is the Hamiltonian, $\hat{N}$ is the particle number operator, and $\hat{I}_m$ is an operator giving the net current in lead m. (We will discuss this operator further below.) These constraints are conveniently imposed using Lagrangian multipliers. The probability entering Eq. (2) can be written as the matrix element $P_\gamma = \langle \gamma | \hat{\rho} | \gamma \rangle$ of the density matrix $\hat{\rho}$. Then maximizing S_I subject to the constraints gives the density matrix

$$\hat{\rho} = \exp\left[-\beta \left(\hat{H} - \mu \hat{N} - \sum_m \xi_m \hat{I}_m \right) \right]. \tag{3}$$

As in ordinary equilibrium thermodynamics, the Lagrangian multiplier associated with the constraint on N is μ, the global chemical potential. The intensive variables ξ_m are Lagrangian multipliers associated with the constraints on the currents. Notice that because of current conservation there are only $M-1$ independent current constraints, and hence one ξ_m can be chosen freely. It turns out to be convenient to choose $\xi_d = 0$, and we do so henceforth. Associated with the constraint on U is the Lagrangian multiplier β. In equilibrium β is the inverse temperature; we shall shortly present several arguments why this continues to be true here.

For clarity let us begin by considering the case of an ideal two-terminal device. This can be either a quasi-one-dimensional conductor, or an ideal Hall bar. We will discuss the latter, since the former then follows easily. In the two-terminal case we can drop the terminal index m, and understand that $k > 0$ ($k < 0$) corresponds to states

injected by the source (drain). These can also be called right and left movers, respectively. In this two-terminal case the density matrix Eq. (3) becomes

$$\hat{\rho} = \exp\left[-\beta(\hat{H} - \mu\hat{N} - \xi\hat{I})\right], \tag{4}$$

where I measures the current carried from source to drain. Suppose now that we have noninteracting electrons (a restriction which can be removed) in an ideal two-terminal, two-dimensional strip lying in the xy plane, subject to a magnetic field along z. Let x be the longitudinal and y the transverse coordinate. The electrons are confined to the strip by a potential $V(y)$ which, for an ideal device, we assume to be a function of the transverse coordinate only. In the Landau gauge $\mathbf{A}(\mathbf{r}) = By\hat{\mathbf{x}}$, the single-particle Hamiltonian is (for charge e electrons of effective mass m^*)

$$\begin{aligned} H &= \frac{1}{2m^*}\left(\mathbf{p} - \frac{e}{c}\mathbf{A}\right)^2 + V(y) \\ &= -\frac{\hbar^2}{2m^*}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right) - \frac{e\hbar By}{im^*c}\frac{\partial}{\partial x} \\ &\quad + \frac{e^2B^2}{2m^*c^2}y^2 + V(y). \end{aligned} \tag{5}$$

In a nonideal two-terminal system this describes the leads in the asymptotic region, i.e., away from the scattering center; there the scattering states $\psi^+_{nk}(\mathbf{r})$ have the asymptotic form[18]

$$\psi^+_{nk}(\mathbf{r}) = \frac{1}{\sqrt{2\pi}} \times \begin{cases} e^{ikx}f_{nk}(y) + \sum\limits_{n'k}{}' r(n'k', nk)\sqrt{\dfrac{k}{k'}}e^{-ik'_{n'}x}f_{n'k'}(y), & x \to \infty \\ \sum\limits_{n'k}{}' t(n'k', nk)\sqrt{\dfrac{k}{k'}}e^{ik'x}f_{n'k'}(y), & x \to -\infty, \end{cases} \tag{6}$$

with $t(n'k', nk)$ and $r(n'k', nk)$ the transmission and reflection amplitudes, respectively. The functions $f_{nk}(y)$ satisfy the Schrödinger equation

$$\left(\frac{\hbar^2k^2}{2m^*} - \frac{\hbar k}{m^*}\frac{eBy}{c} + \frac{1}{2m^*}\frac{e^2B^2y^2}{c^2} + V(y)\right) f_{nk}(y) = \epsilon_{nk}f_{nk}(y) \tag{7}$$

and we have adopted the convention that primed summations are restricted to energy-conserving processes, i.e., $\epsilon_{nk} = \epsilon_{n'k'}$ in Eq. (6).

In this paper we will consider only the case of ideal systems, for which $r(n'k', nk) = 0$ and $t(n'k', nk) = \delta_{n',n}\delta_{k',k}$. In this case instead of an infinite system we may choose a finite length L, and impose periodic boundary conditions. This replaces the scattering states by $\psi_{nk}(\mathbf{r}) = e^{ikx}f_{nk}(y)/\sqrt{L}$. For an ideal system using the periodic eigenstates gives the same results as the scattering states, but the presentation is simpler.

Next we need to construct the current operator $\hat{I}$ which enters Eq. (4). In first-quantized form, the current density operator can be written

$$\mathbf{j}(\mathbf{r}) = \frac{e}{2m^*}\sum_i \left[\mathbf{\Pi}_i\delta(\mathbf{r} - \mathbf{r}_i) + \delta(\mathbf{r} - \mathbf{r}_i)\mathbf{\Pi}_i\right]. \tag{8}$$

Here i labels the particles and $\mathbf{\Pi}_i = \mathbf{p}_i - e\mathbf{A}_i/c$. The total current passing at x is a result of integrating across the strip:

$$\begin{aligned} I(x) &= \int dy j_x(x, y) \\ &= \frac{e}{2m^*}\sum_i \left[\Pi_{ix}\delta(x - x_i) + \delta(x - x_i)\Pi_{ix}\right]. \end{aligned} \tag{9}$$

In general, matrix elements of $I(x)$ between different single-particle states have oscillatory components $\propto \exp[i(k - k')x]$. Such oscillatory components in the diagonal elements do not correspond to any *net* current—the current carried by a scattering state $\psi^+_{nk}(\mathbf{r})$ in the absence of a magnetic field is $e\hbar k/(2\pi m)\tilde{T}_{nk}$, where $\tilde{T}_{nk}$ is the total transmission probability for that state. It is therefore useful to define a current operator which is the average of $I(x)$ over the length of the device:

$$I = \frac{1}{L}\int dx I(x) = \frac{e}{m^*L}\sum_i \Pi_{ix}. \tag{10}$$

In the two-terminal case we can use this as the current operator in Eq. (4). The second quantized form of this is

$$\hat{I} \equiv \frac{1}{L}\sum_{nk}\frac{e\hbar k}{m^*}\hat{c}^\dagger_{nk}\hat{c}_{nk} - \frac{e^2B}{m^*cL}\sum_{nn'k} y_{n'nk}\hat{c}^\dagger_{n'k}\hat{c}_{nk}, \tag{11}$$

where

$$y_{n'nk} \equiv \int dy\, f^*_{n'k}(y) y f_{nk}(y). \tag{12}$$

The second-quantized form is more convenient for an open system, and for an ideal system we can go from a closed to an open system by making the replacement $\frac{1}{L}\sum_k \to \frac{1}{2\pi}\int dk$.

In the absence of an external magnetic field, $\hat{H}$ and $\hat{I}$ commute, so the density matrix is diagonal in the diagonal representation of $\hat{H}$. Hence the thermal occupancy of the single-particle state labeled by n, k is given by

$$f_{nk}(\xi) = \frac{1}{\exp[\beta(\epsilon_{nk} - \xi i_{nk} - \mu)] + 1}. \tag{13}$$

Here $i_{nk} = e\hbar k/m^*L$ is the current of state nk.

In the presence of an external magnetic field, however, $\hat{H}$ and $\hat{I}$ do not commute (despite our suggestion to the contrary[1]), not even for an ideal system. This immediately leads to the apparent problem that the density matrix given by Eq. (3) is not stationary, even though the system is in steady state. One possible solution was given by Grandy,[12] who suggested that, in this case, one should include in the constraints only the parts of the operators which are diagonal in the representation of $\hat{H}$. Grandy shows that the diagonal part $\hat{F}_d$ of an operator $\hat{F}_H$ at time $t = 0$ can in the Heisenberg representation be written as

$$\hat{F}_d = \lim_{\eta\to 0} \eta \int_{-\infty}^{0} \hat{F}_H(t) e^{\eta t}\, dt, \tag{14}$$

where η is a mathematical infinitesimal. The density matrix so constructed is then manifestly time invariant since it commutes with $\hat{H}$.

One could object that this procedure is somewhat *ad hoc*. We believe that also it is not quite correct, and that the correct way to impose a steady-state constraint in the maximum entropy approach is somewhat subtle in the presence of a magnetic field. The heart of the problem lies in how the density matrix Eq. (3) can be made stationary, i.e., how one can find a basis in which both the Hamiltonian $\hat{H} - \mu\hat{N}$ and the generalized canonical Hamiltonian $\hat{H} - \xi\hat{I} - \mu\hat{N}$ are diagonal. We will construct such a basis by following the procedure of adiabatic turning on, a physically appealing approach. That is, we will turn aside from the maximum entropy language of constraints for a while, use adiabatic turning on, and then show how the result is connected to Grandy's idea.

We begin by noting that, in first quantization, $H - \xi I$ can be written as a sum over single-particle terms of the form

$$\frac{1}{2m^*}\mathbf{\Pi}^2 + V - \xi\frac{e}{m^*L}\Pi_x = \frac{1}{2m^*}\left[\left(\mathbf{\Pi} - \frac{e\xi}{L}\hat{\mathbf{x}}\right)^2 - \left(\frac{e\xi}{L}\right)^2\right] + V \tag{15}$$

[using Eqs. (5) and (10)]. This expression is, within an additive constant, formally identical to adding a pure gauge vector potential $(c\xi/L)\hat{\mathbf{x}}$ to the original Hamiltonian. Now suppose that $c\xi$ is an integral number of flux quanta (a condition that can be obtained as closely as desired for a long system). In this case the vector potential $(c\xi/L)\hat{\mathbf{x}}$ can be removed by a trivial gauge transformation. From this we conclude that the spectra of $H - \xi I$ and H are identical (within an additive constant).

This motivates an important step: let us make the physical interpretation that the steady state is reached by adiabatically turning on such an extra vector potential. Do this by switching on a vector potential $\mathbf{A}_\xi(t) = -[c\xi(t)/L]\hat{\mathbf{x}}$, where $\xi(t) = \xi g(t)$ and $g(t)$ is a dimensionless function satisfying $g(t \to -\infty) = 0$, $g(t = 0) = 1$, and $g'(t = 0) = 0$. (This generates the current adiabatically via an electric field $\mathbf{E} = -(1/c)\partial\mathbf{A}/\partial t$ which, for the periodic system, is caused by the adiabatic insertion of an integral number of magnetic flux quanta through the center of the system.) The time-dependent single-particle Hamiltonian under this adiabatic switching on is

$$H(t) = \frac{1}{2m^*}\left(\mathbf{\Pi} - \frac{e}{c}\mathbf{A}_\xi(t)\right)^2 + V(y). \tag{16}$$

The Hamiltonian at $t = 0$ differs from that at $t = -\infty$ by a pure gauge transformation:

$$H(0) = e^{iS} H(-\infty) e^{-iS}, \tag{17}$$

where

$$S = -\tilde{\xi}x \quad \text{and} \quad \tilde{\xi} = e\xi/\hbar L. \tag{18}$$

Let $U(t_2, t_1)$ be the unitary time evolution operator (in the Schrödinger representation) corresponding to $H(t)$, including the adiabatic evolution of $\xi(t)$. Then an initial $t = -\infty$ single-particle eigenstate $|nk, -\infty\rangle$ will evolve at $t = 0$ to $|nk, 0\rangle = U(0, -\infty)|nk, -\infty\rangle$. This $t = 0$ state is an eigenstate of $H(0)$. Because of Eq. (17), it is also the gauge transformation of some other eigenstate of $H(-\infty)$:

$$|nk, 0\rangle = e^{iS}|n, k + \tilde{\xi}, -\infty\rangle. \tag{19}$$

Equation (19) can be shown by noting that (i) the spectra of $H(0)$ and $H(-\infty)$ are identical; (ii) k is a good quantum number at both $t = -\infty$ and $t = 0$, and only the state with wave vector $k + \tilde{\xi}$ on the right-hand side gives k on the left; and (iii) only one subband can have the correct energy and wave vector.

To proceed, let us assume that the system reaches equilibrium (at $t = -\infty$) before the current is switched on, and that the system is then thermally isolated while the current is turned on adiabatically. (This is the common assumption of nonequilibrium statistical mechanics.[25]) Then the final electronic distribution can be described by an evolved $t = 0$ density matrix which here, because of the equality Eq. (19), can be written in a particularly simple form using the $t = -\infty$ states. To do this let us change to the language of many-particle states. Define an inital $t = -\infty$ N-electron state

$$|\alpha KN; -\infty\rangle = \prod_{\text{occ}} \hat{c}^\dagger_{nk}(\text{vacuum}), \tag{20}$$

for which

$$\hat{H}(-\infty)|\alpha KN; -\infty\rangle = E_{\alpha KN}|\alpha KN; -\infty\rangle. \tag{21}$$

Here $K = \sum_{\text{occ}} k$, and "occ" signifies a certain set of occupied single-particle states (index α distinguishes among the different many-particle states with momentum K). The time evolution is governed by $\hat{H}(t)$, which contains the adiabatically changing $\xi(t)$, and the corresponding time evolution operator $\hat{U}$: $|\alpha KN; t\rangle = \hat{U}(t, -\infty)|\alpha KN; -\infty\rangle$. Because $\hat{H}(0) = e^{i\hat{S}}\hat{H}(-\infty)e^{-i\hat{S}}$,

$$\hat{H}(0)|\alpha KN; 0\rangle = E_{\alpha, K+N\tilde{\xi}, N}|\alpha KN; 0\rangle. \tag{22}$$

The expectation value of any operator $\hat{O}$ in the Schrödinger picture at $t = 0$ can be written

$$\mathrm{Tr}\left(\hat{O}(t=0)\hat{\rho}(t=0)\right) = \sum_{\alpha KN} \langle \alpha KN;0|\hat{O}|\alpha KN;0\rangle P_{\alpha KN;0}. \quad (23)$$

Here $P_{\alpha NK;t} = \langle \alpha KN;t|\hat{\rho}(t)|\alpha KN;t\rangle$ are expectation values of the density matrix at time t. The density matrix at $t = 0$ is presumed to be determined by the (equilibrium) density matrix at $t = -\infty$. Hence its $t = 0$ expectation values are simply $P_{\alpha KN;0} = P_{\alpha KN;-\infty} = e^{-\beta(E_{\alpha KN}-\mu N)}$. Again using the gauge transformation between $\hat{H}(-\infty)$ and $\hat{H}(0)$, we can thus write Eq. (23) as

$$\begin{aligned}\mathrm{Tr}\left(\hat{O}(t=0)\hat{\rho}(t=0)\right) &= \sum_{\alpha KN} \langle \alpha KN;0|\hat{O}|\alpha KN;0\rangle e^{-\beta(E_{\alpha KN}-\mu N)} \\ &= \sum_{\alpha KN} \langle \alpha, K+\tilde{\xi}N, N;-\infty|e^{-i\hat{S}}\hat{O}e^{i\hat{S}}|\alpha, K+\tilde{\xi}N, N;-\infty\rangle e^{-\beta(E_{\alpha KN}-\mu N)} \\ &= \sum_{\alpha KN} \langle \alpha KN;-\infty|e^{-i\hat{S}}\hat{O}e^{i\hat{S}}|\alpha KN;-\infty\rangle e^{-\beta(E_{\alpha,K-\tilde{\xi}N,N}-\mu N)}. \end{aligned} \quad (24)$$

Here $e^{-i\hat{S}}\hat{O}e^{i\hat{S}}$ is simply the operator $\hat{O}$ in the gauge in which ξ has been removed. Hence Eq. (24) shows that we can write thermal averages in a very simple way by tracing over the eigenstates of the original unevolved Hamiltonian, that is, the Hamiltonian without the current-inducing term $\xi(t)$. The third form in Eq. (24) is the most convenient for calculations in the presence of a magnetic field.

Now let us make one final step which will connect the preceding to the maximum entropy approach. Rewrite the first of Eqs. (24) using $E_{\alpha KN} = E_{\alpha,K+N\tilde{\xi},N} - (E_{\alpha,K+N\tilde{\xi},N} - E_{\alpha,KN})$ and

$$\begin{aligned}E_{\alpha,K+N\tilde{\xi},N} - E_{\alpha KN} &= \int_0^{\tilde{\xi}} \frac{\partial E_{\alpha,K+N\tilde{\xi}',N}}{\partial \tilde{\xi}'} d\tilde{\xi}' \\ &= \tilde{\xi}\int_{-\infty}^{0} \frac{\partial E_{\alpha,K+N\tilde{\xi}',N}}{\partial \tilde{\xi}'} g'(t)\, dt\ , \end{aligned} \quad (25)$$

where we have put $\tilde{\xi}' = \tilde{\xi}g(t)$ and $g'(t) = dg/dt$. In the Schrödinger representation,

$$\begin{aligned}\frac{\partial E_{\alpha,K+N\tilde{\xi}',N}}{\partial \tilde{\xi}'} &= \frac{\partial}{\partial \tilde{\xi}'}\langle \alpha KN;t|\hat{H}(t)|\alpha KN;t\rangle \\ &= \frac{\hbar L}{e}\langle \alpha KN;t|\hat{I}(t)|\alpha KN;t\rangle. \end{aligned} \quad (26)$$

The second step follows immediately by inserting the time-evolution operator, since the contributions from the derivatives of this operator with respect to $\tilde{\xi}'$ cancel out in this case. Therefore we can write the thermal averages in the form

$$\begin{aligned}\mathrm{Tr}\left(\hat{O}(t=0)\hat{\rho}(t=0)\right) &= \sum_{\alpha KN} \langle \alpha KN;-\infty|\hat{O}_H(0)|\alpha KN;-\infty\rangle \\ &\quad \times \langle \alpha KN;-\infty|\hat{\rho}(t=0)|\alpha KN;-\infty\rangle, \end{aligned} \quad (27)$$

where the subscript H means the Heisenberg representation $[\hat{O}_H(t) \equiv \hat{U}^\dagger(t,-\infty)\hat{O}(t)\hat{U}(t,-\infty)]$ and where

$$\hat{\rho}(t=0) = \exp\left[-\beta\left(\hat{H}_H(0) - \xi\int_{-\infty}^{0} \hat{I}_H(t)g'(t)dt - \mu\hat{N}_H(0)\right)\right]. \quad (28)$$

This density matrix is diagonal in the diagonal representation of the operator in its exponent. This expression is similar to that given by Grandy [Eq. (14)], and as in that case is stationary in this steady-state case. Thus the density matrix Eq. (28) can be obtained using the maximum entropy approach, following a steady-state prescription quite similar to Grandy's. However, we note that the adiabatic turning on of the vector potential $\mathbf{A}_\xi$ enters into the dynamics of the system in a nontrivial way: in Eq. (28) the time dependence is given by the Hamiltonian $\hat{H}(t)$ in which the adiabatic turning on is included. This apparently technical point is very important to get the correct answer.

Notice also that we can formally expand $E_{\alpha,K-\tilde{\xi}N,N}$ in the third of Eqs. (24) in a Taylor series to obtain

$$E_{\alpha,K-\tilde{\xi}N,N} = E_{\alpha KN} - \tilde{\xi}N\frac{\partial E_{\alpha KN}}{\partial K} - \frac{1}{2!}\tilde{\xi}N^2\frac{\partial^2 E_{\alpha KN}}{\partial K^2} + \cdots. \quad (29)$$

For the case of a parabolic confining potential, $V(y) = \frac{1}{2}m^*\omega_0^2 y^2$, the single-particle eigenvalues are quadratic in k, and the expansion terminates at the term quadratic in ξ:

$$E_{\alpha,K-\tilde{\xi}N,N} = E_{\alpha KN} - \xi I_{\alpha KN} - N\frac{e^2\xi^2}{2m^*L^2}\frac{1}{1+\omega_c^2/\omega_0^2}. \quad (30)$$

Thus the density matrix for a quadratic confining poten-

tial and a magnetic field has the same form as that of a system *without* a magnetic field, and we can in both cases write the single-particle occupancies as in Eq. (13).

In summary, the occupancies we obtain depend on the current carried by a state as well as its energy. In the absence of a magnetic field, or with a parabolic confining potential, the occupancies for noninteracting electrons take the simple form given by Eq. (13). In the presence of a magnetic field, it is most convenient to use the form shown in the third of Eqs. (24). For noninteracting electrons this corresponds to occupancies

$$f_{nk}(\xi) = \frac{1}{\exp[\beta(\epsilon_{n,k-\tilde{\xi}} - \mu)] + 1}. \tag{31}$$

This reduces to Eq. (13) in the absence of a magnetic field, or for the case of a parabolic confining potential. By Eq. (28) this can be seen to be a generalization of the simple (zero-magnetic-field) result in which the steady-state condition is carefully enforced.

These nonequilibrium distributions in the presence of a current are simply a shift of the equilibrium occupancies, which slide over an amount $\tilde{\xi}$ in k space. With only one subband ($n = 0$) occupied, the single-particle distributions we obtain are similar to those of the LB approach. For example, with zero magnetic field, at zero temperature, states are occupied up to an energy $\mu+\xi i_{0k}(\xi)$ (with k here the highest occupied state on an edge). This is pictured in Fig. 3(a). Because the current has opposite signs for positive and negative k, right- and left-moving states are occupied up to different energies. The combination $\mu + \xi i_{0k}$ acts in this case like an effective local chemical potential. In the general case with several subbands occupied, however, states are occupied up to different energies in each subband. [See Fig. 3(b).] At zero temperature the states in subband n are occupied up to $\mu + \xi i_{nk}$ which depends on the current carried by the highest occupied state in this subband. This current is in general different in different subbands, and so the distributions f_{nk} cannot be described in terms of local chemical potentials. Said another way, if one insisted on defining local chemical potentials, there would need to be a different "local chemical potential" for each subband.

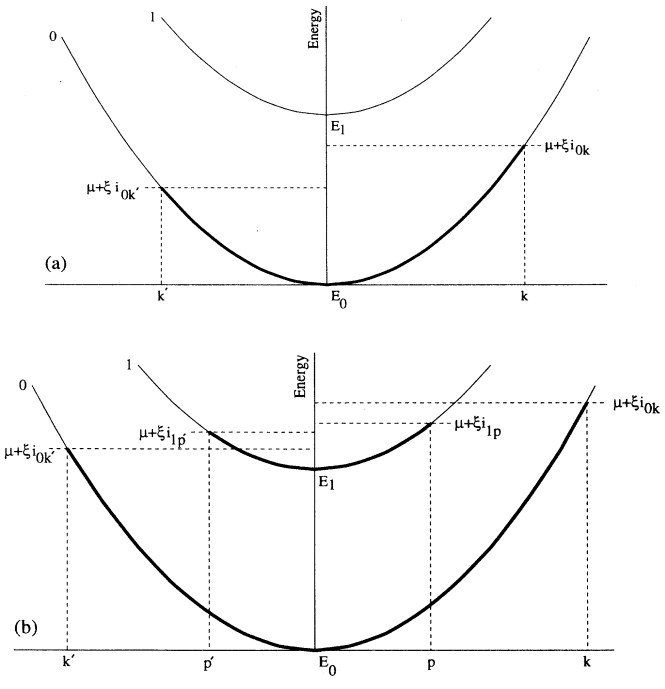

FIG. 3. Current-dependent occupancies of states in an ideal two-terminal device, from the maximum entropy method. Occupied states (at zero temperature) are indicated by a heavy line. (a) When only one subband is occupied, states are occupied up to energies $\mu + \xi i_{0k}$. These have different values for the outermost occupied states with $k > 0$ and $k < 0$ (shown at k and k', respectively). (b) When more than one subband is occupied, states in different subbands are occupied up to different energies. Compare this with Fig. 1.

The generalization of these results to the multiterminal case is very straightforward and is presented in the Appendix. In the absence of a magnetic field, or for a parabolic confinement, the result is very simple. In the multiterminal system the occupancy of the scattering states $|\psi^{+}_{mnk}\rangle$ (incoming in terminal m) is given, for noninteracting electrons, by

$$f_{mnk}(\xi) = \frac{1}{\exp\left[\beta\left(\epsilon_{mnk} - \sum_{m'} \xi_{m'} i_{m',mnk} - \mu\right)\right] + 1} \tag{32}$$

[cf. the corresponding two-terminal expression Eq. (13)].

B. Calculating voltages

Within the reservoir model it is quite clear that ordinary voltage measurements at a contact correspond to the local chemical potential: $V_m = \mu_m/e$. (See, for example, Ref. 16.) As we have emphasized above, the distributions derived from the MEA cannot in general be described in terms of a local chemical potential, and clearly the LB prescription for finding V_m cannot apply. Nor does the MEA itself give some procedure for determining voltages. The approach we take to calculate voltages comes from the following physical picture: a voltmeter determines the voltage differences between two terminals by measuring the work required to move a small amount of charge from one to the other. If moving some charge δQ takes a work δW, then the voltage difference is the ratio $\delta W/\delta Q$. The problem is then to calculate this work.

The work required to move reversibly between equilibrium states is given by changes in thermodynamic potentials (e.g., the Helmholtz free energy, when the temperature is constant). In general potentials cannot be defined in nonequilibrium systems, which ordinarily involve dissipation. The problem of steady-state transport in mesoscopic systems, however, is a special case of nonequilibrium thermodynamics. In the device there is no inelastic scattering; this, and the steady-state condition, permit us to define thermodynamic potentials of

the electron distribution.[26] For example, the information entropy is equivalent to an ordinary thermodynamic entropy, as mentioned above and explained below; and so here we can define the Helmholtz free energy as usual, $F = U - TS$. The work δW done on the system at constant temperature is then equal to the change in free energy,

$$\delta F = \mu\,\delta N + \sum_m \xi_m\,\delta I_m. \tag{33}$$

That is, at constant temperature the free energy is a function of the electron number and of the net current at each terminal. However, the variables most easily varied in the theory are the intensive variables $\mu, \xi_s, \xi_1, \ldots, \xi_{M-2}$. Varying any of these in general changes the occupancy f_α of incoming states at each terminal, by Eq. (13). Here for definiteness we will use the representation given by the scattering states $|\psi^+_{mnk}\rangle$ labeled by $\alpha = mnk$, with occupancies f_{mnk}. In what follows it is useful to define a quantity N_m which is the number of occupied scattering states entering at terminal m. Thus $N_m = \sum_{nk} f_{mnk}$ and the total electron number in the device is $N = \sum_m N_m$.

Let η refer to any of the independent variables $\mu, \xi_s, \xi_1, \ldots, \xi_{M-2}$. Let δI^η_m denote the change in the net current entering at terminal m when variable η is varied while the other independent variables are held fixed. Similarly let δN^η_m be the corresponding change in N_m. Then changing $\eta \to \eta + \delta\eta$ produces a free energy change

$$\delta F^\eta = \sum_m \left(\mu \delta N^\eta_m + \xi_m\,\delta I^\eta_m\right). \tag{34}$$

We obtain the potentials V_m at the terminals by interpreting this free energy change as the work done to add δN^η_m electrons against the voltage V_m at the terminals. At each terminal the energy cost is $e\delta N^\eta_m V_m$. Thus we can also write

$$\delta F^\eta = \sum_m e\,\delta N^\eta_m V_m. \tag{35}$$

Equating Eqs. (34) and (35) for each of the M variables η gives a set of M linearly independent equations,

$$\sum_m \delta N^\eta_m (eV_m - \mu) = \sum_m \xi_m \delta I^\eta_m, \tag{36}$$

which are to be solved for the unknown terminal voltages V_m. The nonlocal resistance measured between two terminals m and m' is then $R_{sd,mm'} = (V_m - V_{m'})/I$. The matrix δN^η_m on the left-hand side of Eq. (36) is invertible, and the resulting potentials are then automatically given relative to the global chemical potential μ.

C. Example

In this section we will illustrate the approach outlined above by finding the resistance of an ideal two-terminal system, and then extend the result to the ideal multiterminal case. First consider a two-terminal device. For this we can use eigenstates which satisfy periodic boundary conditions on a length L along the device. This particular choice of boundary condition gives a simple density of states, but is of no other significance. The states are labeled by a subband index n and wave vector k. Again in this case we can drop the terminal subscripts and understand that states with $k > 0$ are injected by the source and $k < 0$ by the drain. We choose $\xi_d = 0$ and write ξ for ξ_s. The net current I_s entering at the source is just the total current I through the device. Here we consider only one simple case, for which the states have energy $\epsilon_{nk} = E_n + \hbar^2 k^2/2m^*$ and current $i_{nk} = e\hbar k/m^* L$. This can represent one-dimensional (1D) transport, or a Hall bar with a parabolic transverse confinement. (In the latter case, $\hbar^2/2m^* \to \omega_0 \ell^2/2$, where ω_0 is the curvature of the confining potential and $\ell = \sqrt{\hbar c/eB}$ is the magnetic length.) By Eq. (30), the occupancy f_{nk} depends on energy and current in the combination $\epsilon_{nk} - \xi i_{nk}$. This is

$$E_n + \frac{\hbar^2 k^2}{2m^*} - \xi\frac{e\hbar k}{m^* L} = E_n + \frac{\hbar^2}{2m^*}\left(k - \frac{e\xi}{\hbar L}\right)^2 - \frac{e^2\xi^2}{2m^* L^2}. \tag{37}$$

Thus the occupancy can be written

$$f_{nk} = f(\tilde\epsilon_{nk} - \tilde\mu) \equiv 1/(e^{\beta(\tilde\epsilon_{nk} - \tilde\mu)} + 1), \tag{38}$$

where we have defined

$$\tilde\epsilon_{nk} = E_n + \frac{\hbar^2}{2m^*}(k - \tilde\xi)^2, \tag{39a}$$

$$\tilde\xi = \frac{e\xi}{\hbar L}, \tag{39b}$$

$$\tilde\mu = \mu + \frac{\hbar^2\tilde\xi^2}{2m^*}, \tag{39c}$$

for reasons which will shortly be apparent. The total number of electrons occupying current-carrying states in the device is

$$N = \sum_{nk} f(\tilde\epsilon_{nk} - \tilde\mu) = \sum_n \int_{E_n}^\infty d\tilde\epsilon\,\rho_n(\tilde\epsilon) f(\tilde\epsilon - \tilde\mu), \tag{40}$$

where in the limit of large L we can take the k's and hence $\tilde\epsilon_{nk}$ to be quasicontinuous, and define for the latter a 1D density of states,

$$\rho_n(\tilde\epsilon) = (L/\pi)[2\hbar^2(\tilde\epsilon - E_n)/m^*]^{-1/2}. \tag{41}$$

From the integral form in Eq. (40) we see that N depends on μ and ξ only in the combination $\tilde\mu$ given by Eq. (39c). Thus here $\tilde\mu$ acts like a current-dependent chemical potential controlling the number of occupied current-carrying states. The total current carried by the occupied states is

$$I = \sum_{nk} i_{nk} f(\tilde\epsilon_{nk} - \tilde\mu). \tag{42}$$

Notice that f_{nk} is symmetric about $k = \tilde\xi$. That is, the

two states in subband n at $k_\pm = \tilde{\xi} \pm [2m^*(\tilde{\epsilon} - E_n)/\hbar^2]^{1/2}$ have the same value of $\tilde{\epsilon}_{nk} = \tilde{\epsilon}$ and hence, from Eq. (38), the same occupancy. The current carried by these two states is then

$$i_n(k_+) + i_n(k_-) = e\hbar(k_+ + k_-)/m^*L = 2e\hbar\tilde{\xi}/m^*L. \tag{43}$$

Thus the total current is

$$I = \frac{e\hbar\tilde{\xi}}{m^*L}\sum_{nk} f(\tilde{\epsilon}_{nk} - \tilde{\mu}) = \frac{e\hbar}{m^*L}\tilde{\xi}N. \tag{44}$$

With these expressions it is now straightforward to compute the current-voltage relations by inverting Eqs. (36). In this two-terminal example, it is easier to use $\tilde{\mu}$ and $\tilde{\xi}$ as independent variables instead of the original pair μ, ξ. Then for this example there are two Eqs. (36), labeled by $\eta = \tilde{\mu}, \tilde{\xi}$:

$$\begin{pmatrix} \delta N_s^{\tilde{\mu}} & \delta N_d^{\tilde{\mu}} \\ \delta N_s^{\tilde{\xi}} & \delta N_d^{\tilde{\xi}} \end{pmatrix} \begin{pmatrix} eV_s - \mu \\ eV_d - \mu \end{pmatrix} = \xi \begin{pmatrix} \delta I^{\tilde{\mu}} \\ \delta I^{\tilde{\xi}} \end{pmatrix}. \tag{45}$$

These equations contain δN_m^η, the change in the occupation number N_m of states which carry current into terminal m. We are interested in the changes in these which result when $\eta = \tilde{\xi}, \tilde{\mu}$ are changed infinitesimally. Rather than computing these directly, consider a function

$$Q(\tilde{\mu}, \tilde{\xi}; k_0) = \sum_{nk} f(\tilde{\epsilon}_{nk} - \tilde{\mu})\mathrm{sgn}(k - k_0). \tag{46}$$

Notice that Q depends on $\tilde{\xi}$ via $\tilde{\epsilon}_{nk}$, which, by Eq. (39a), depends on $\tilde{\xi}$. The quantity $\frac{1}{2}(N \pm Q)$ gives the number of occupied states with $k \gtrless k_0$. We ultimately seek the changes in the number of occupied states with $k \gtrless 0$ due to infinitesimal changes $\delta\tilde{\mu}$ and $\delta\tilde{\xi}$. But the changes in occupation of these states are the same as the changes in occupation numbers of states with $k \gtrless k_0$, as long as k_0 is not too close to the edge of occupied k's (that is, as long as states near wave number k_0 are either fully occupied or unoccupied). Thus we can write

$$\delta N_{s,d}^\eta = \frac{1}{2}(\delta N^\eta \pm \delta Q^\eta), \tag{47}$$

where the upper (lower) sign is for $m = s$ (d), and η represents variations in either $\tilde{\mu}$ or $\tilde{\xi}$. Let us choose $k_0 = \tilde{\xi}$. [Then, more precisely, Eq. (47) is true as long as the region of k space where $f(\tilde{\epsilon}_{nk} - \tilde{\mu})$ is very close to unity (at low temperatures) brackets both $\tilde{\xi}$ and $k = 0$. This will be discussed further in part II, since this point leads to a breakdown in quantization in point contacts.] Clearly $Q(\tilde{\mu}, \tilde{\xi}; k_0 = \tilde{\xi}) = 0$. We seek

$$\begin{aligned} \delta Q &\equiv Q(\tilde{\mu} + \delta\tilde{\mu}, \tilde{\xi} + \delta\tilde{\xi}; \tilde{\xi}) \\ &= \sum_{nk} \frac{1}{e^{\beta(\tilde{\epsilon}'_{nk} - \tilde{\mu} - \delta\tilde{\mu})} + 1}\mathrm{sgn}(k - \tilde{\xi}), \end{aligned} \tag{48}$$

where $\tilde{\epsilon}'_{nk} = E_n + \hbar^2(k - \tilde{\xi} - \delta\tilde{\xi})^2/2m^*$. Since $\tilde{\epsilon}'_{nk}$ is symmetric about $\tilde{\xi} + \delta\tilde{\xi}$, the only contribution to the sum over k is for $\tilde{\xi} \le k \le \tilde{\xi} + 2\delta\tilde{\xi}$, so that

$$\begin{aligned} \delta Q &= \sum_n \sum_{k=\tilde{\xi}}^{\tilde{\xi}+2\delta\tilde{\xi}} \left[e^{\beta(\tilde{\epsilon}'_{nk} - \tilde{\mu} - \delta\tilde{\mu})} + 1\right]^{-1} \\ &\approx \frac{L}{\pi}\delta\tilde{\xi}\sum_n \left[e^{\beta(E_n - \tilde{\mu})} + 1\right]^{-1} \end{aligned} \tag{49}$$

plus terms of second order in the quantities $\delta\tilde{\xi}$ and $\delta\tilde{\mu}$. That is, $\delta Q^{\tilde{\xi}} = \delta Q$ and $\delta Q^{\tilde{\mu}} = 0$. From the expression for N given in Eq. (40) it is evident that varying $\tilde{\xi}$ and $\tilde{\mu}$ changes N by

$$\delta N = -\delta\tilde{\mu}\sum_n \int_{E_n}^\infty d\tilde{\epsilon}\rho_n(\tilde{\epsilon})\frac{\partial f(\tilde{\epsilon} - \tilde{\mu})}{\partial\tilde{\epsilon}}, \tag{50}$$

that is, $\delta N^{\tilde{\mu}} = \delta N$ and $\delta N^{\tilde{\xi}} = 0$. Combining Eqs. (50), (49), and (47), we find

$$\begin{aligned} \delta N_s^{\tilde{\mu}} &= \delta N_d^{\tilde{\mu}} = \frac{1}{2}\delta N, \\ \delta N_s^{\tilde{\xi}} &= -\delta N_d^{\tilde{\xi}} = \frac{1}{2}\delta Q. \end{aligned} \tag{51}$$

From Eq. (44) we obtain

$$\begin{aligned} \delta I^{\tilde{\mu}} &= \frac{e\hbar}{m^*L}\tilde{\xi}\delta N, \\ \delta I^{\tilde{\xi}} &= \frac{e\hbar}{m^*L}N\delta\tilde{\xi}. \end{aligned} \tag{52}$$

It is now a simple matter to invert the matrix in Eq. (45) and solve for $V_{s,d}$. We write the final answer as the two-terminal conductance

$$G = I/(V_s - V_d) = \frac{e^2}{h}\sum_n \frac{1}{e^{\beta(E_n - \tilde{\mu})} + 1}. \tag{53}$$

If $\tilde{\mu}$ exceeds the band minima of the first j subbands (or Landau levels), then at low temperatures $G = je^2/h$. Finite-temperature corrections are exponentially small. This quantization occurs because at large currents it is quite possible for states in one subband to be occupied up to energies above those of empty states in other subbands [as pictured in Fig. 3(a)]—not because relaxation has failed to occur, but because according to the MEA this is the steady-state (although highly nonequilibrium) result.

This exact result is true for an ideal system with a parabolic energy spectrum. We emphasize that this example is special only in that it can be solved analytically (which is the reason we have used it to illustrate our method). We have numerically studied nonparabolic energies ϵ_{nk} and multiterminal systems, and find in these cases that the accuracy of quantization is limited only by

the numerical accuracy. In our method zero-temperature quantization persists up to very high currents and voltages. Quantization can fail when the current grows so large that a subband is occupied only for carriers moving in one direction. This is similar to the distributions for which the LB approach fails to give quantization [as in Fig. 1(b)]. But it is important to emphasize that the conditions under which this failure of quantization occurs are very different in the two cases. In our approach, increasing the current shifts the occupancy of states in a subband (at zero temperature) from left movers to right movers. In IQHE devices one should expect the subbands to look schematically like those in Figs. 1 and 3: a relatively flat "bulk" region surrounded by edge regions where the energy increases rapidly. As long as the density of bulk states greatly exceeds that of edge states, one can push the energy of the highest-occupied right movers very high while left movers in the same subband are occupied. In the LB approach, on the other hand, a subband immediately becomes partially occupied as soon as μ_s but not μ_d exceeds the subband's minimum [as in Fig. 1(b)]. This occurs whenever eV_H exceeds the subband spacing. In our approach, instead, it is quite possible [as in Fig. 3(a)] for states in a lower subband to be occupied to energies above vacant states in a higher subband. To empty the left movers in a subband would generally require quite high currents and voltages in our approach. How high the voltage can get depends on the ratio of the number of current-carrying states in the relatively flat portion of the band to those in the steep edge region. This should be quite high in typical Hall devices. In the particular case of a parabolic energy spectrum, however, left movers would be drained at relatively low voltages (eV_H of order $2\hbar\omega_c$). This is not relevant to Hall devices, but is relevant to experiments in quasi-one-dimensional systems such as quantum point contacts, and to the saturation observed there.[24] This will be discussed further in part II. Exceptions to quantization can also occur in our approach when the current is large enough to induce breakdown in the sample as a consequence of other dissipative mechanisms.[27] For ordinary samples exhibiting the IQHE, neither this nor the earlier possibility occurs, and our approach described above provides an explanation of the extremely accurate quantization observed outside the linear response regime. This ability to explain the high-precision IQHE experiments is nontrivial, and lends credence to the postulate of the MEA.

The above example was explained in detail to show how our approach is applied in general. In this particular case of a two-terminal system, a simplification is possible. Since N depends on $\tilde{\mu}$ and not $\tilde{\xi}$, changing the latter while holding the former constant corresponds to moving a charge $e\delta Q/2$ from one terminal to the other (or from edge to edge in a quantum Hall sample). The work required to do this (at constant N) is $\delta F^{\tilde{\xi}} = \xi \delta I^{\tilde{\xi}}$. Hence the voltage difference is

$$e(V_s - V_d) = \frac{\delta F^{\tilde{\xi}}}{e\delta Q/2}, \tag{54}$$

which gives the same result.

III. DISCUSSION

A. Density matrix

The density matrix and distribution which result from the maximum entropy approach, and which lie at the heart of our ability to explain the high-precision IQHE experiments, are unusual, but not unheard of in the literature. The distribution f_α in Eq. (13) was proposed earlier by Heinonen and Taylor,[28] and more recently by Ng.[29] These authors argued that in a device without dissipation it should be possible to define a free energy which, presumably, would be minimized. The process of minimization, subject to the current constraint, is formally identical to the MEA's maximization of information entropy, and leads to the same distribution. In this paper we have obtained this result on considerably more fundamental grounds—assuming only the postulates of the maximum entropy approach to nonequilibrium statistical mechanics are, in fact, correct.

Here we will seek some understanding of the density matrix Eq. (3) and distribution Eq. (13) from other viewpoints. Note first of all that this density matrix has the general form which Hershfield recently showed should exist on quite general grounds in steady-state nonequilibrium systems.[30] In his rather more conventional approach to nonequilibrium statistical mechanics, it is quite evident that β in Eq. (13) is indeed $1/k_BT$, where T is the temperature. This was not obvious in the MEA, but based on Hershfield's work we can make the same identification, in the case of steady-state mesoscopic transport. It also then follows that the information and thermodynamic entropies are equivalent [with c in Eq. (2) set to k_B].

The distributions Eq. (13) are obviously quite different from the LB distributions. The latter follow in a very straightforward way from the model of terminals as reservoirs, and are widely considered valid by their ability to describe nonlocal resistances in many low-current experiments, as well as by the appealing simplicity and clarity of the reservoir model itself. How can this difference in distributions be understood physically? The LB distributions can be derived from linear response theory, with a certain assumption about how the system is driven. One can apply something very similar to the reservoir idea in a very precise calculation by modeling the leads as ideal and supposing that far from the device there is in each lead a well-defined electrochemical potential. Stone and co-workers have shown that this assumption leads *in linear response* to the multiterminal LB formalism.[7,8] We should note, however, that this work has been criticized by Landauer, based on their use of leads of constant cross section, which do not have the geometrical spreading he believes is necessary for the reservoir picture.[10] Even so, while indeed no large reservoir is invoked in the work of Stone *et al.*, a very similar idea—that far from the device a probe can be described as a system at constant potential—is at the heart of this calculation, entering as a boundary condition. (We should also note that, since in principle linear response can be calculated using equi-

librium distributions, the success of the linear theory in predicting low-current properties does not imply that the linear distributions are correct.)

We emphasize that built into Stone *et al.*'s and related calculations is the model of terminals as entities described by local chemical potentials. This model fits neatly into the most common way to approach nonequilibrium problems: assuming that there are two large reservoirs, each in equilibrium, and that transport (say) between them occurs when they are connected by a small channel. This is a very familiar approach to nonequilibrium statistical mechanics. And yet it amounts to an assumption as to how the system is driven. In typical transport experiments in the IQHE, say, a constant current source is connected to the sample. Using the reservoir picture amounts to a model of how a current source (when connected to a mesoscopic device) actually drives the current. Perhaps it is valid in linear response to model the current source as two leads at different potentials. Even if valid in linear response, it is not *a priori* clear that the picture should be valid at large currents. In fact, the failure of the LB approach at large currents suggests not.

Let us examine this in more detail. Suppose that the current source itself can be viewed as an object which gives rise to different potentials in the physical leads. Then the current is carried down long macroscopic lengths of lead until it is injected into the device through the terminals. Perhaps along the macroscopic wire the local potential changes gradually and smoothly (due to elastic and inelastic scattering); that is, perhaps the distribution locally can be described by a local chemical potential. Does this picture work right up to the vicinity of the mesoscopic device? In fact, there are several reasons to think not.

The simplest way to approach this is in the approximation of the Boltzmann approach to transport.[31] Consider an ordinary dissipative conductor carrying a low current. Let us suppose that the current is driven by reservoirs held at two different potentials. Then at the reservoirs the Boltzmann distribution $f(\mathbf{r},\mathbf{k},t)$ will have the LB form. But a perfectly standard Boltzmann calculation shows that far from the reservoirs the distribution evolves into a current-dependent form identical (in first order) to that derived from the MEA [Eq. (13)], and different from the first-order LB distributions. Suppose that the ends of an effectively one-dimensional conductor (at $x = \pm L/2$) are held at different electrochemical potentials $\mu \pm \Delta\mu/2$. The distribution is then labeled by a wave number k plus a subband index n: $f = f(x,n,k,t)$. Suppose further that this gives rise to a uniform electric field $E = \Delta\mu/eL$ through the conductor. Then in steady state and one dimension the Boltzmann equation becomes[31]

$$v\frac{\partial f}{\partial x} + \frac{\partial k}{\partial t}\frac{\partial f}{\partial k} = \left(\frac{df}{dt}\right)_{\text{coll}}. \tag{55}$$

We want the lowest-order (in $\Delta\mu$) solution to this, in the relaxation time approximation.[31] Since the conductor's ends are held at definite local chemical potentials, the distribution takes the LB form at the ends:

$$f(\mp L/2,n,k,t) = 1/(\exp\{\beta[\epsilon_{nk} - (\mu \pm \Delta\mu/2)]\} + 1). \tag{56}$$

The upper sign is for positive k and the lower for negative k. Equation (55) is a completely standard Boltzmann problem, with the only wrinkle provided by the boundary conditions, Eq. (56). To linear order this problem is solved by

$$f = f_0 + \Delta\mu\frac{\partial f_0}{\partial\epsilon}\left[\left(\frac{\hbar k\tau}{mL} \mp \frac{1}{2}\right)e^{-m(x\pm L/2)/\hbar k\tau} - \frac{\hbar k\tau}{mL}\right], \tag{57}$$

where again the upper (lower) sign is for $k > 0$ ($k < 0$). Here $f_0(\epsilon_{nk}) = 1/(e^{\beta(\epsilon_{nk}-\mu)} + 1)$ is the equilibrium distribution and τ is the relaxation time (which can depend on k). Let us suppose that the relaxation time τ is much less than the transit time across the system: $\tau \ll L/v$, where the speed $v = \hbar k/m$. Then at distances much farther than $v\tau$ from the ends the distribution becomes

$$f = f_0 - \frac{\tau\Delta\mu}{e} i_{nk}\frac{\partial f_0}{\partial\epsilon}. \tag{58}$$

Here $i_{nk} = ev/L$ is the current carried by the state in subband n with wave vector k. Equation (58) is the lowest-order term in an expansion in current of

$$f_0(\epsilon_{nk} - \xi i_{nk}) = \frac{1}{\exp[\beta(\epsilon_{nk} - \mu - \xi i_{nk})] + 1}, \tag{59}$$

where we have here identified $\xi = \tau\Delta\mu/e$. Note that this is not equal to the corresponding first-order LB distribution $f = f_0 \mp (\Delta\mu/2)\partial f_0/\partial\epsilon$.

That is, in a completely typical dissipative conductor, the Boltzmann distribution is precisely the lowest-order approximation to the distribution [Eq. (13)] obtained from the maximum entropy calculation. This is the case even though we chose the boundaries to model reservoirs; the distribution evolves from the LB form near the ends to the current-dependent form away from the ends. Our point with this example is the following. Even though the LB distribution has the authority of widespread usage, perhaps one should instead typically expect to find current-dependent distributions carrying steady-state currents into a mesoscopic system; for the dissipative wires carrying the current to the mesoscopic device should themselves typically have such distributions. Note that the Boltzmann equation gives current-dependent distributions also in the familiar case of a three-dimensional conductor in the relaxation time approximation. In this case the resulting distribution is obtained by displacing the Fermi surface in the direction of the current, and therefore here too the occupancy of single-particle states depends on their current as well as their energy.

There are special cases in which current-dependent distributions can be obtained in other ways. We have argued elsewhere that this should arise by Galilean transforma-

tion in a translationally invariant dissipationless system.[1] We also have given elsewhere a detailed example in which a mechanism to switch on the current is provided, and which results in this distribution.[32] These, and the general result of Hershfield mentioned earlier,[30] all support our identification of β as $1/k_BT$, as well as the notion that these distributions should be generally expected in steady-state transport.

If the reservoir picture is only adequate at low currents, how can one picture the way a current source drives a large current through a mesoscopic system? Perhaps one can think of the current source as forcing current through a region full of scatterers, like someone being forced to run a gauntlet. The current source pushes electrons in; they scatter into other states. In steady state it is reasonable that the occupancies of the various states will be influenced by their ability to carry current.

Finally, we have emphasized that the distributions resulting from the MEA (given the constraints of particle number, internal energy, and current) cannot be described in terms of a local chemical potential. Nonetheless it is evident from our calculation of the terminal voltages V_m that the quantity eV_m plays the role of a kind of a local chemical potential. That is, eV_m is equal to the energy cost required to add an extra particle to the terminal (more precisely, to occupy an extra incoming state at terminal m). This is clearly not a local chemical potential in the LB sense—the MEA distributions in terminal m are not of the Fermi-Dirac form with a local chemical potential eV_m.

B. Maximum entropy approach

Perhaps the most unorthodox part of our calculation is its use of the maximum entropy approach. The essential feature of the MEA (besides the obvious fact of the entropy maximization) is that observables enter the formalism as constraints. For example, here we have treated the current source—the object driving the system out of equilibrium—merely as something which imposes a constraint on the total current I. That is, the result of the driving enters the formalism. In a typical linear response calculation, the driver enters as a term in the Hamiltonian (say, a small electric field). A difficulty of the MEA is that it provides no prescription for how one should determine the constraints. It seems that one must be guided by the physical picture. This has been called "the basic problem" with the maximum entropy approach.[33] In many cases it simply is not possible to know what the constraints are—for example, in the case of hot electrons in semiconductors it appears that one must somehow incorporate information about phonon interactions as a constraint;[33] and nobody knows how to do this. But the fact that this approach may be difficult in some problems does not of course mean that it is always difficult. In fact, steady-state mesoscopic transport seems to be ideally suited for this approach. Guided by the physical picture, we have made the simplest possible supposition about constraints, and it appears to work. In particular, since the current source is designed to hold the current constant, we simply treat the current as a constraint.

A second difficulty with the MEA is that often it is difficult to calculate what the microstates are which enter the formalism. In the case of steady-state mesoscopic transport, this is not a problem. Since the thermodynamic system is the mesoscopic device plus ideal leads (in the scattering geometry), it is quite straightforward to calculate the entire set of microstates. Here we have emphasized how to do this for noninteracting electrons, but it is also possible in the interacting case. We will discuss this more in part II.

It is interesting to note that the LB distribution can also be obtained from the MEA by a different choice of constraint. This happens if one assumes that the current source somehow constrains the particle numbers N_m entering at each terminal, rather than determining the net currents at each terminal.[29] In the MEA these constraints lead to Lagrangian multipliers μ_m, and the resulting occupancies are the LB distributions $f^{\rm LB}_{mnk}$. Thus one might be tempted to ascribe the difference between the LB distributions and those obtained by us to the way in which the current source is modeled. At low currents the use of local chemical potentials can be justified using linear response theory, viewing the potential difference as driving the current. This cannot be extended to high currents. Since the LB distribution is associated with an ordinary electrochemical potential at each reservoir, one might suppose that it models a voltage source instead of a current source. If so, then the I-V curve at large currents and voltages would depend on whether voltage or (as is usual) current is applied.[29] (In the linear regime, both approaches give the same result.) In fact, based on our arguments in Sec. III A this appears unlikely. Even if a voltage source is applied to the ends of the macroscopic wires leading to a device, it appears that by the time one moves far from the source (i.e., gets near the device) one should expect the distribution to have evolved to a current-dependent form.

Finally, we note that our formalism provides a natural connection between Laughlin's original[14] gauge-invariance argument for the IQHE in closed systems and the more recent explanations of the IQHE in multiterminal *open* systems based on edge states in the LB formalism.[16] While the connection between edge states and bulk response is well known for a closed system, there has been no apparent way to connect the LB formalism with its central concepts of reservoirs injecting scattering states with the gauge arguments. Our formalism does provide such a connection directly between gauge fields and scattering states. The connection is made by interpreting the Lagrangian parameters ξ_m, which arise from the current constraint in the MEA, as vector potentials which are turned on adiabatically.

ACKNOWLEDGMENTS

We gratefully acknowledge discussions with C. T. Van Degrift, E. Palm, S. Girvin, S. Hershfield, M. Büttiker, and A. D. Stone. This work was supported in part by

the UCF Division of Sponsored Research, and by the National Science Foundation under Grant No. DMR-9301433.

APPENDIX: DISTRIBUTIONS FOR MULTITERMINAL SYSTEMS

It is straightforward to generalize our formalism to open ideal multiterminal devices in the absence of an external magnetic field. We define coordinates (x_m, y_m) such that $\hat{\mathbf{x}}_m$ points into the device in lead m, and $\hat{\mathbf{x}}_m \times \hat{\mathbf{y}}_m = \hat{\mathbf{z}}$; the confining potential $V(\mathbf{r})$ (which defines the geometry) can asymptotically in lead m be written $V(y_m)$. In an ideal system, all states injected into lead m reach only the terminal at the lead m', say, and we define the operator for incoming current in lead m'' of states injected into lead m as

$$I_{m'';m} = \frac{\hbar e}{2\pi m^*} \sum_n \int dk\, k \hat{c}^\dagger_{mnk} \hat{c}_{mnk}, \quad m'' = m, \quad \text{(A1)}$$

and

$$I_{m'';m} = -\frac{\hbar e}{2\pi m^*} \sum_n \int dk\, k \hat{c}^\dagger_{mnk} \hat{c}_{mnk}, \quad m'' = m' , \quad \text{(A2)}$$

and zero otherwise. Just as in the two-terminal case, the density matrix then commutes with the Hamiltonian $\hat{H}$ and one can immediately write down the single-particle occupation numbers in terms of the diagonal elements ϵ_{mnk} and i_{mnk}.

For a multiterminal device in the presence of an external magnetic field, we start by choosing a gauge in which the vector potential is a Landau gauge in all leads,[8]

$$\mathbf{A}(\mathbf{r}) = B y_m \hat{\mathbf{x}}_m , \quad \text{(A3)}$$

with x_m in the asymptotic region in lead m. States injected into lead m now all exit the system at the next lead m' defined by the direction of the magnetic field. For example, in Fig. 2, if the source is terminal m, then terminal 1 in that figure would correspond to m', and so on. We define the current operator for current in lead m'' of states injected into lead m as

$$I_{m'';m} = \frac{\hbar e}{2\pi m^*} \sum_n \int dk\, k \hat{c}^\dagger_{mnk} \hat{c}_{mnk} - \frac{e^2 B}{2\pi m^* c} \sum_{nn'} \int dk\, k y_{m,n',n,k} \hat{c}^\dagger_{mn'k} \hat{c}_{mnk}, \quad m'' = m, \quad \text{(A4)}$$

and

$$I_{m'';m} = -\frac{\hbar e}{2\pi m^*} \sum_n \int dk\, k \hat{c}^\dagger_{mnk} \hat{c}_{mnk} + \frac{e^2 B}{2\pi m^* c} \sum_{nn'} \int dk\, k y_{m,n',n,k} \hat{c}^\dagger_{mn'k} \hat{c}_{mnk}, \quad m'' = m' \quad \text{(A5)}$$

and zero otherwise. The generalized canonical Hamiltonian $\hat{H} - \sum_m \xi_m \hat{I}_m - \mu \hat{N}$, which determines the density matrix, is in the asymptotic regions in lead m equivalent to a Hamiltonian given in first quantization by

$$H = \frac{1}{2m^*} \sum_i \left[\mathbf{\Pi}_i - \hbar \widetilde{\xi}_m \hat{\mathbf{x}}_m \right]^2 - \frac{\hbar^2 \widetilde{\xi}_m^2}{2m^*} + V(y_m), \quad \text{(A6)}$$

with

$$\widetilde{\xi}_m = \frac{e \xi_m}{2\pi\hbar}. \quad \text{(A7)}$$

Thus $\widetilde{\xi}_m$ enters as a vector potential in lead m. We then smoothly continue $\widetilde{\xi}_m \hat{\mathbf{x}}_m$ to a divergenceless irrotational vector function $\widetilde{\mathbf{\Xi}}_m(\mathbf{r})$ in all space which is equal to $\widetilde{\xi}_m \hat{\mathbf{x}}_m$ asymptotically in lead m, $-\widetilde{\xi}_m \hat{\mathbf{x}}_{m'}$ asymptotically in lead m' (as defined above), and zero in the asymptotic regions in all other leads. A suitable analogy would be to think of $\widetilde{\mathbf{\Xi}}_m(\mathbf{r})$ as the velocity field of an incompressible fluid which flows irrotationally in through lead m and out through lead m'. We then turn on the vector potentials adiabatically by taking $\xi_m \to \xi_m g(t)$ and study the time-dependent system with the Hamiltonian

$$\widetilde{H} = \frac{1}{2m^*} \sum_i \left\{ \left[\mathbf{\Pi}_i - \frac{e}{c} \sum_m \mathbf{A}_{\xi_m}(t, \mathbf{r}_i) \right]^2 + V(\mathbf{r}_i) \right\}, \quad \text{(A8)}$$

where $\mathbf{A}_{\xi_m}(t, \mathbf{r}) = -(c\hbar/e) \widetilde{\mathbf{\Xi}}_m(t, \mathbf{r})$. Just as in the ideal two-terminal case, adiabatically turning on the vector potentials from $t = -\infty$ will then generate a system at $t = 0$ with net currents. Here, the parameters ξ_m have to be thought of as adjustable parameters, adjusted to give the correct net currents at each terminal. As an example, consider an ideal multiterminal IQHE system. Suppose there are four identical terminals, as shown in Fig. 2. Then all of the current leaving the source flows along the lower edge to terminal 1. For the net current through terminal 1 to vanish, it is necessary that $\xi_1 = \xi_s$. Similarly, $\xi_2 = \xi_d = 0$. The work required to transfer a certain charge from terminal 1 to 2 is then precisely that required to move the same charge from s to d in the two-terminal calculation. Hence the two-terminal conductance calculated above becomes here the Hall conductance, which is then quantized.

[1] O. Heinonen and M.D. Johnson, Phys. Rev. Lett. **71**, 1447 (1993).
[2] K. von Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980); K. von Klitzing, Rev. Mod. Phys. **58**, 519 (1986); see also *The Quantum Hall Effect*, edited by R.E. Prange and S.M. Girvin (Springer-Verlag, New York, 1987).
[3] For example, the GaAs heterojunction used to maintain the resistance standard at the U.S. National Institute of Standards and Technology shows quantization to better than one part per billion at the $\nu = 4$ plateau for $eV_H > 16\hbar\omega_c$.
[4] See, for example, J.J. Palacios and C. Tejedor, Phys. Rev. B **44**, 8157 (1991).
[5] R. Landauer, IBM J. Res. Dev. **1**, 223 (1957); Philos. Mag **21**, 863 (1970); Z. Phys. B **21**, 247 (1975); **68**, 217 (1987).
[6] M. Büttiker, Phys. Rev. Lett. **57**, 1761 (1986); Phys. Rev. B **38**, 9375 (1988); IBM J. Res. Dev. **32**, 317 (1988).
[7] A.D. Stone and A. Szafer, IBM J. Res. Dev. **32**, 384 (1988).
[8] H.U. Baranger and A.D. Stone, Phys. Rev. B **40**, 8169 (1989).
[9] P.C. van Son and T.M. Klapwijk, Europhys. Lett. **12**, 429 (1990).
[10] R. Landauer, in *Analogies in Optics and Micro Electronics*, edited by W. van Haeringen and D. Lenstra (Kluwer, Dordrecht, 1990), p. 243; Phys. Scr. **T42**, 110 (1992).
[11] See, for example, *Quantum Transport in Semiconductors*, edited by D.K. Ferry and C. Jacoboni (Plenum, New York, 1992).
[12] E.T. Jaynes, Phys. Rev. **106**, 620 (1957); **108**, 171 (1957). For recent treatments, see, for example, W.T. Grandy, Jr., in *Foundations of Statistical Mechanics* (Reidel, Dordrecht, 1988), Vols. 1 and 2; H.S. Robertson, *Statistical Thermophysics* (Prentice-Hall, Englewood Cliffs, NJ, 1993).
[13] Electron-electron interactions may be included in general.
[14] R.B. Laughlin, Phys. Rev. B **23**, 5632 (1981).
[15] Y. Imry, in *Directions in Condensed Matter Physics*, edited by G. Grinstein and G. Mazenko (World Scientific, Singapore, 1986), p. 101.
[16] For a review of the Landauer-Büttiker approach in the context of the quantum Hall effect, see M. Büttiker, in *Nanostructured Systems* (Ref. 34), Chap. 4, p. 191.
[17] F. Sols, Ann. Phys. (N.Y.) **214**, 386 (1992).
[18] A.M. Kriman, N.C. Kluksdahl, and D.K. Ferry, Phys. Rev. B **36**, 5953 (1987).
[19] Note that in this somewhat unfortunate, but standard, notation the field creation operator is written $\hat{\psi}^\dagger(\mathbf{r}) = \sum_{mnk} \left(\psi^+_{mnk}\right)^\dagger (\mathbf{r}) \hat{c}^\dagger_{mnk}$.
[20] P. Streda, J. Kucera, and A.H. MacDonald, Phys. Rev. Lett. **59**, 1973 (1987); J.K. Jain and S.A. Kivelson, *ibid.* **60**, 1542 (1988); Phys. Rev. B **37**, 4276 (1988).
[21] M. Büttiker, in *Feskörperprobleme. Advances in Solid State Physics*, edited by F. V. Braunschweig (Pergamon, Oxford, 1990), Vol. 30, p. 41.
[22] U. Sivan and Y. Imry, Phys. Rev. B **33**, 551 (1986).
[23] B.I. Halperin, Phys. Rev. B **25**, 2185 (1982).
[24] H. van Houten, C.W.J. Beenakker, and B.J. van Wees, in *Nanostructured Systems* (Ref. 34).
[25] For a clear statement about this point, see J. Rammer and H. Smith, Rev. Mod. Phys. **58**, 323 (1986).
[26] More general circumstances under which thermodynamic potentials can be defined out of equilibrium have been considered by R.J. Tykodi, *Thermodynamics of Steady State* (Macmillan, New York, 1967).
[27] M. Cage *et al.*, Phys. Rev. Lett. **51**, 1374 (1983).
[28] O. Heinonen and P.L. Taylor, Phys. Rev. B **28**, 6119 (1982); **32**, 633 (1985).
[29] T.K. Ng, Phys. Rev. Lett. **68**, 1018 (1992).
[30] S. Hershfield, Phys. Rev. Lett. **70**, 2134 (1993).
[31] G.D. Mahan, *Many-Particle Physics*, 2nd ed. (Plenum, New York, 1990), Chap. 7, p. 602.
[32] O. Heinonen and M.D. Johnson, Phys. Rev. B **49**, 13 740 (1994).
[33] R. Landauer, Physica A **194**, 551 (1993).
[34] *Nanostructured Systems*, edited by M. Reed, Semiconductors and Semimetals Vol. 35 (Academic Press, New York, 1992).