The noncommutative geometry of the quantum Hall effect

J. Bellissard,[a] A. van Elst,[b] and H. Schulz- Baldes[c]
Université Paul Sabatier, Toulouse, France[d]

(Received 1 May 1994; accepted for publication 17 May 1994)

An overview of the integer quantum Hall effect is given. A mathematical framework using nonommutative geometry as defined by Connes is prepared. Within this framework, it is proved that the Hall conductivity is quantized and that plateaux occur when the Fermi energy varies in a region of localized states.

TABLE OF CONTENTS

0022-2488/94/35(10)/5373/79/$6.00
J. Math. Phys. **35** (10), October 1994

I. INTRODUCTION

In 1880, Hall[1] undertook the classical experiment which led to the so-called Hall effect. A century later, von Klitzing and his co-workers[2] showed that the Hall conductivity was quantized at very low temperatures as an integer multiple of the universal constant e^2/h. Here e is the electron charge whereas h is Planck's constant. This is the integer quantum Hall effect (IQHE). For this discovery, which led to a new accurate measurement of the fine structure constant and a new definition of the standard of resistance,[3] von Klitzing was awarded the Nobel price in 1985.

On the other hand, during the seventies, Connes[4,5] extended most of the tools of differential geometry to noncommutative C^*-algebras, thus creating a new branch of mathematics called *Noncommutative Geometry*. The main new result obtained in this field was the definition of cyclic cohomology and the proof of an index theorem for elliptic operators on a foliated manifold. For this work and also his contribution to the study of von Neumann algebras, Connes was awarded the Fields Medal in 1982. He recently extended this theory to what is now called *Quantum Calculus*.[6]

After the works by Laughlin[7] and especially by Kohmoto, den Nijs, Nightingale, and Thouless[8] (called TKN_2 below), it became clear that the quantization of the Hall conductance at low temperature had a geometric origin. The universality of this effect had then an explanation. Moreover, as proposed by Prange,[9,10] Thouless,[11] and Halperin,[51] the plateaux of the Hall conductance which appear while changing the magnetic field or the charge-carrier density, are due to localization. Neither the original Laughlin paper nor the TKN_2 one however could give a description of both properties in the same model. Developing a mathematical framework able to reconcile topological and localization properties at once was a challenging problem. Attempts were made by Avron *et al.*[12] who exhibited quantization but were not able to prove that these quantum numbers were insensitive to disorder. In 1986, Kunz[13] went further on and managed to prove this for disorder small enough to avoid filling the gaps between Landau levels.

But in Refs. 14–16 one of us proposed to use noncommutative geometry to extend the TKN_2 argument to the case of arbitrary magnetic field and disordered crystal. It turned out that the condition under which plateaux occur was precisely the finiteness of the localization length near the Fermi level. This work was rephrased later on by Avron *et al.*[17] in terms of charge transport and relative index, filling the remaining gap between experimental observations, theoretical intuition and mathematical frame.

Our aim in this work is to review these various contributions in a synthetic and detailed way. We will use this opportunity to give proofs that are missing or scattered in the literature. In addition, we will discuss the effect of disorder from two complementary aspects.

On the one hand, we will develop our point of view on localization produced by quenched disorder. This is crucial for understanding the IQHE. We review various localization criteria and formulate them in terms of noncommutative geometry. With the Dixmier trace, Connes introduced a remarkable technique into quantum calculus. In our context, it allows us to give the precise condition under which the Hall conductance is quantized; this condition is shown to be a localization condition.

On the other hand, we also propose a model for electronic transport giving rise to the so-called "relaxation time approximation" and allowing to derive a Kubo formula for the conductivity. This approach allows us to describe the effect of time-dependent disorder in a phenomenological way. This latter has quite different consequences from those of the quenched disorder such as a nonzero finite direct conductivity. Even though this approach is not original in its principle, the noncommutative framework allows us to treat the case of aperiodic crystals and magnetic fields when Bloch theory fails. Therefore, strictly speaking, our Kubo formula is new. We also show, without proofs, how to justify the linear response theory within this framework, leaving the formal proofs for a future work. The advantage of this approach is to give control of the various approximations that have to be made to fit the ideal result with experiments. For this reason, we discuss the effects of temperature, of nonlinear terms in the electric field, of the finite size of samples and finally

those of collisions and disorder. In particular, we argue that the discrepancy $\delta\sigma_H$ between the measured Hall conductivity and the ideal one, given by a Chern number, is dominated by the collision terms. In the center of a plateau, we get the rough estimate

$$\frac{\delta\sigma_H}{\sigma_H} \leqslant \text{const } \nu\frac{e}{h}\frac{\lambda^2}{\mu_c}, \tag{1}$$

where ν is the filling factor, λ is the localization length, and μ_c is the charge-carrier mobility. e/h is a universal constant, ν about unity and the localization length typically of the order of the magnetic length. Inserting measured values for the mobility, one obtains 10^{-4} for the right-hand side expression. This estimate does not take into account the Mott conductivity. However, it shows why both a large quenched disorder (in order to have small localization lengths) and a large mobility (namely, a low collision rate) are necessary in order to get accurate measurements. Such a compromise is realized in heterojunctions and to less extent in MOSFETs. The estimate (1) also permits us to understand intuitively why the plateaux in the fractional quantum Hall effect (FQHE) are less precise, since the localization length of Laughlin quasiparticles is probably larger than that of electrons at integer plateaux, and their mobility probably lower.

No attempt will be made however to extend our noncommutative approach to the FQHE. We will only give some insight and a short review of works that we feel relevant in view of a mathematically complete description of the FQHE.

This rest of the article is organized as follows. In Sec. II we give the conventional explanations of the IQHE. In particular, we discuss the Laughlin argument, the topological aspect introduced by TKN_2 and the effects of localization in a qualitative way. Section III is devoted to the mathematical framework needed for noncommutative geometry. In particular we describe how to overcome the difficulty of not having Bloch's theorem for aperiodic media. We then show that the Brillouin zone still exists as a noncommutative manifold. We also give the main steps of our strategy leading to a complete mathematical description of the IQHE. In Sec. IV we discuss transport theory leading to Kubo's formula. We show that in the IQHE idealization, the Hall conductance is a noncommutative Chern number. We also relate this Chern number to a Fredholm index which leads to the quantization of the Hall conductance. Through the notion of relative index we show in which sense this approach is a rigorous version of the Laughlin argument. Section V is devoted to localization theory. We give various criteria and define various localization lengths which are commonly used in the literature. We also show how to express these notions in the noncommutative language. This part allows us to explain on a rigorous basis the occurrence of plateaux of the Hall conductance. Finally, we show that such criteria are in fact satisfied in models such as the Anderson model. In Sec. VI we give some consequences of this theory for practical models. In particular we show that low-lying states do not contribute to the IQHE. We also discuss the open question where the jumps of the Hall conductance occur. Section VII is a short review of available results on the FQHE.

II. IQHE: EXPERIMENTS AND THEORIES

A. The classical Hall effect

Let us consider a very flat conductor placed in a constant uniform magnetic field in the z direction perpendicular to the plane Oxy of the plate (see Fig. 1). If we force a constant current in the x direction, the electron fluid will be submitted to the Lorentz force perpendicular to the current and the magnetic field. Hall realized that the electron fluid is incompressible so that the Lorentz force must produce a pressure, namely, a potential difference perpendicular to the current.

Let $\mathbf{j}$ be the current density, $\mathcal{B}$ the magnetic field, and $\mathcal{E}$ the Hall electric field. In a stationary state, the electric forces acting on the charges are opposite to the Lorentz forces. This leads to the equation

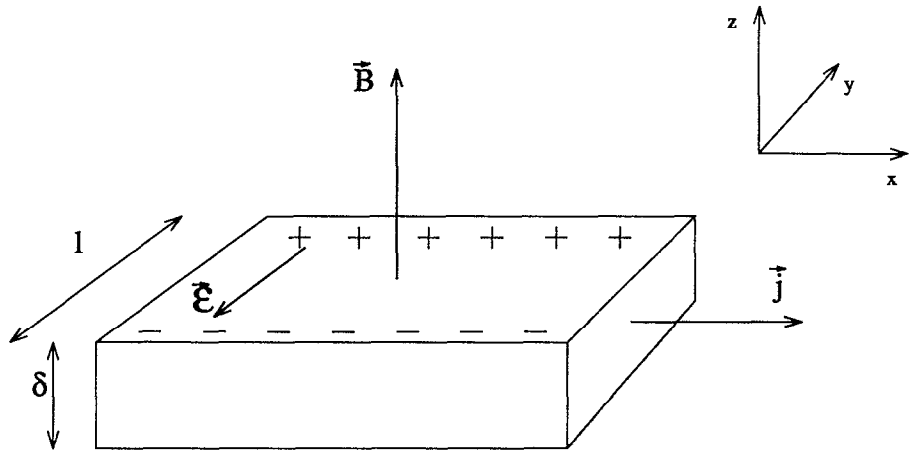

FIG. 1. The classical Hall effect: the sample is a thin metallic plate of width δ. The magnetic field $\mathscr{B}$ is uniform and perpendicular to the plate. The current density $\mathbf{j}$ parallel to the x-axis is stationary. The magnetic field pushes the charges as indicated creating the electric field $\mathscr{E}$ along the y direction. The Hall voltage is measured between opposite sides along the y-axis.

$$nq\mathscr{E}+\mathbf{j}\times\mathscr{B}=0, \tag{2}$$

where n is the charge-carrier density and q is the charge of the carriers. Since the magnetic field $\mathscr{B}$ is perpendicular to both $\mathbf{j}$ and $\mathscr{E}$, solving (2) for $\mathbf{j}$ gives

$$\mathbf{j}=\frac{nq}{\mathscr{B}^2}\,\mathscr{B}\times\mathscr{E}=\sigma\mathscr{E},$$

where $\mathscr{B}$ is the modulus of the magnetic field and σ is the conductivity tensor. The antidiagonal components of the tensor are the only nonvanishing ones and can be written as $\pm\sigma_H\delta$, where δ is the plate width and σ_H is called the Hall conductance. Thus

$$\sigma_H=\frac{qn\delta}{\mathscr{B}}.$$

We remark that the sign of σ_H depends upon the sign of the carrier charge. In particular, the orientation of the Hall field will change when passing from electrons to holes. Both possibilities were already observed by Hall using various metals. This observation is commonly used nowadays to determine which kind of particles carries the current.

Let ℓ be the plate width in the y direction (see Fig. 1). The current intensity inside the plate is then given by $I=j\delta\ell$ where j is the modulus of $\mathbf{j}$. The potential difference created by the Hall field is $V_H=-\ell\mathscr{E}.\mathbf{u}$ if $\mathbf{u}$ is the unit vector along the y axis. Using (2) we find

$$V_H=\frac{\mathscr{B}I}{nq\delta}=\frac{I}{\sigma_H}. \tag{3}$$

In particular, for a given current intensity I, the thinner the plate the higher the potential difference. For example, for a good conductor like gold at room temperature, the charge carrier density is of order of $6\times10^{28}m^{-3}$ (see Ref. 18, Chap. 1). Thus, for a magnetic field of 1 T, a current intensity of 1 A and a potential difference of 1 mV the plate width is about 1 μm. These numbers

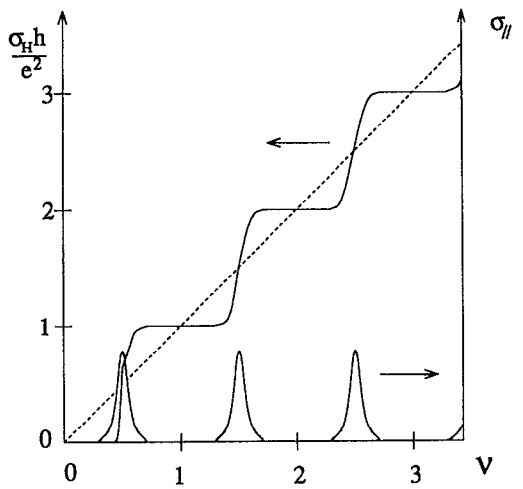

FIG. 2. Schematic representation of the experimental observations in the IQHE. The Hall conductivity σ_H is drawn in units of e^2/h vs filling factor ν. The dashed line shows the Hall conductivity of the Landau Hamiltonian without disorder. The direct conductivity $\sigma_{\parallel}$ is shown in arbitrary units.

explain why the effect was so difficult to observe. It forced Hall to use very thin gold leaves in the beginning. In modern devices, much thinner "plates" with thickness of about 100 Å are produced in inversion layers between two semiconductors.

In view of (3), the Hall conductance has the dimension of the inverse of a resistance. Since the product $n\delta$ is the number of charge carriers per unit area, the dimensionless ratio

$$\nu=\frac{n\delta h}{\mathscr{B}e},$$

called the *filling factor*, represents the fraction of a Landau level filled by conduction electrons of the thin plate. In terms of this parameter, we obtain for a free electron gas:

$$\sigma_H=\frac{\nu}{R_H}, \quad R_H=\frac{h}{e^2}, \tag{4}$$

where R_H is called the Hall resistance. It is a universal constant with value $R_H=25\,812.80\ \Omega$. R_H can be measured directly with an accuracy better than 10^{-8} in QHE experiments. Since January 1990, this is the new standard of resistance at the national bureau of standards.[3]

B. The quantum Hall effect

Let us concentrate upon the dependence of the Hall conductance (in units of e^2/h) on the filling factor ν. In the classical Hall effect, these two quantities are just equal [Eq. (4)]. Lowering the temperature below 1 K leads to the observation of plateaux for integer values of the Hall conductance (see Fig. 2). In von Klitzing's experiment,[2] the variation of ν was obtained by changing the charge carrier density, whereas in later experiments one preferred varying the magnetic field. The accuracy of the Hall conductance on the plateaux is better than 10^{-8}. For values of the filling factor corresponding to the plateaux, the direct conductivity $\sigma_{//}$, namely the conductivity along the current density axis, vanishes. These two observations are actually the most important ones. The main problems to be explained are the following:

(i) Why do the plateaux appear exactly at integer values?
(ii) How do the plateaux appear?
(iii) Why are these plateaux related to the vanishing of the direct conductivity?

To observe the QHE, physicists have used conduction electrons trapped in the vicinity of an interface between two semiconductors. The local potential difference between the two sides produces a bending of the local Fermi level. Near the interface, this Fermi level meets with the valence band creating states liable to participate in the conductivity. This bending occurs on a distance of the order of 100 Å from the interface, so that the charge carriers are effectively concentrated within such a thin strip. In addition, by changing the potential difference between the two sides, the so-called gate voltage, one can control the charge-carrier density.

The samples used in QHE experiments belong to two different categories. The first one is called MOSFET,[19] for metal-oxide silicon field effect transistors. The interface separates doped silicon from silicon oxide. This device was common in the beginning of the eighties and was the one used in von Klitzing's experiment. However, the electron mobility is relatively low because the control of the flatness of the interface is difficult.

The samples of the other category are heterojunctions. The interface separates GaAs from an alloy of $Al_xGa_{1-x}As$. This kind of device nowadays makes available interfaces almost without any defects. Moreover, electrons therein have a high mobility. These devices are most commonly used in modern quantum Hall experiments.

In both kinds of samples, there are many sources of defects producing microscopic disorder. The first comes from the doping ions. Even though they are usually far from the interface (about 1000 Å), the long range Coulomb potential they produce is strong enough to influence the charges on the interface. It is not possible to control the position of these ions in the crystal. The second source of defects is the roughness of the interface. This is an important effect in MOSFET's, much less in heterojunctions. In the latter the accuracy is better than one atomic layer in every 1000 Å along the interface.[20] Finally, long range density modulation of the compounds may produce visible effects. This is the case especially for heterojunctions where the aluminium concentration may vary by a few percent on a scale of 1 μm.[21]

It is important to notice that the observation of plateaux supposes several conditions.

(i) The effect is more easily seen if the electron fluid is concentrated in such a thin region that it can be considered as two-dimensional. In fact, owing to the trapping effect of the potential interface, the motion perpendicular to the interface is quantized. For good samples in high magnetic fields, the energy difference between two corresponding eigenvalues is big compared to k_BT (where T is the temperature and k_B the Boltzmann constant), so that only the lowest such level has to be considered. Hence the problem becomes effectively two-dimensional.
(ii) The plateaux disappear beyond a temperature of a few Kelvin. More exact, the inelastic relaxation time has to be large enough; otherwise corrections will be needed in formulas calculating the current; this will lead to the destruction of the plateaux. This is the reason why the IQHE is seen more easily in heterojunctions than in MOSFETs: the electron mobility in the former is higher than in the latter.
(iii) We will see that some quenched disorder producing only elastic scattering is necessary for the appearance of the plateaux (see Sec. II D). In practice, the disorder that occurs is strong enough to produce a filling of the gaps between Landau levels.[21,22]
(iv) Clearly the sample size must be big enough as to allow the use of the infinite volume limit. Mesoscopic systems exhibit conductance fluctuations from sample to sample which may partially distroy the effect. Finite volume effects however have been shown to decrease exponentially fast with the sample size.

(v) Finally, the electric field needed to produce the current has to be small. If it is too high, nonlinear phenomena may distroy the plateaux.[23]

Provided the previous conditions are satisfied, the quantum Hall effect is a universal phenomenon. It is quite independent of the specific shape of the sample. Hall plateaux have also been observed in microwave experiments where the topology of the sample is trivial.[24] Note, however, that the centers of the plateaux need not be located near integer values of the filling factor. It depends upon which kind of doping ion is used.[25] The values of the Hall conductance, however, on the plateaux are independent of the nature of the used sample.

C. The Hall effect for the free Fermi gas

Let us first consider a very simple model in which the charge carriers are spinless, free, two-dimensional fermions with charge q. Our aim is to show that no quantization of the Hall conductance is observed in this case.

Since the particles are independent, the quantum motion is described by the one-particle Hamiltonian. Let $\mathbf{A}=(A_1,A_2)$ be the vector potential given by

$$\partial_1 A_2 - \partial_2 A_1 = \mathscr{B},$$

where $\mathscr{B}$ is the modulus of the magnetic field. The energy operator is then given by the Landau Hamiltonian[26]

$$H_L = \frac{(\mathbf{P}-q\mathbf{A})^2}{2m_*}, \tag{5}$$

where $\mathbf{P}$ is the 2D momentum operator and m_* is the effective mass of the particle. This operator is not translation invariant owing to the symmetry breaking produced by the vector potential. However, if one replaces the usual representation of the translation group by the so-called *magnetic translations,*[27] the Landau Hamiltonian becomes translation invariant. Let us introduce the quasimomentum operators $\mathbf{K}=(\mathbf{P}-q\mathbf{A})/\hbar$; they fulfill the canonical commutation relations

$$[K_1,K_2] = \imath \frac{q\mathscr{B}}{\hbar}.$$

Here $q\mathscr{B}/\hbar$ plays the rôle of an effective Planck constant. We see that the Landau Hamiltonian describes a harmonic oscillator so that its spectrum is given by the Landau levels, namely,

$$E_n = \hbar\omega_c(n+\tfrac{1}{2}),$$

where $\omega_c = q\mathscr{B}/m_*$ is the cyclotron frequency and $n \in \mathbf{N}$. Each Landau level is infinitely degenerate owing to translation invariance. The degeneracy per unit area is finite however and given by $q\mathscr{B}/h$. Perturbing this operator will give rise to a band spectrum.

Let us now compute the current. We assume the system to be in a thermodynamical equilibrium. In order to allow nonzero current, we describe the system in the grand-canonical ensemble; in practice such a fluid is open. Since the particles are independent fermions, the thermal averaged density per unit volume of a one-particle translation invariant extensive observable $\mathscr{O}$ at temperature T and chemical potential μ is given by

$$\langle \mathscr{O} \rangle_{T,\mu} = \lim_{\Lambda \uparrow \mathbf{R}^2} \frac{1}{|\Lambda|} \mathrm{Tr}_\Lambda (f_{T,\mu}(H_L)\mathscr{O}),$$

where $f_{T,\mu}(E)=(1+e^{\beta(E-\mu)})^{-1}$ is the Fermi-distribution function and $\beta=1/k_BT$. In the limit above, Λ denotes a square box centered at the origin, and $|\Lambda|$ is its area. The current is represented by the operator

$$\mathbf{J}=\frac{\imath q}{\hbar}[H_L,\mathbf{X}],$$

where $\mathbf{X}$ is the position operator. It is easy to check that this current operator commutes with the magnetic translations. Obviously the thermal average of the current vanishes since no current can flow without an external source of energy. To make it nonzero, we switch on an electric field $\mathscr{E}$ at time $t=0$. For simplicity, we assume this field to be uniform and time-independent. After the field has been switched on, the time evolution is given by

$$\frac{d\mathbf{J}}{dt}=\frac{\imath}{\hbar}[H_{L,\mathscr{E}},\mathbf{J}], \quad \text{with} \quad H_{L,\mathscr{E}}=H_L-q\mathscr{E}\mathbf{X}.$$

The solution of this equation is elementary. Using complexified variables, that is $M=M_1+\imath M_2$ whenever $\mathbf{M}=(M_1,M_2)$ is a 2D vector, we find

$$J(t)=-\imath q\,\frac{\mathscr{E}}{\mathscr{B}}+e^{-\imath\omega_c t}J_0\,,$$

where J_0 is some initial datum. This solution consists of a time-independent part and of an oscillating part with the cyclotron frequency. The time average is just the constant part; it is the system's response to the applied electric field. Taking the thermal average of this constant part, we have the observed current, namely, if $j=j_1+\imath j_2$,

$$j=\lim_{t\to\infty}\int_0^t\frac{ds}{t}\langle J(s)\rangle_{T,\mu}=-\imath\,\frac{q\mathscr{E}}{\mathscr{B}}\lim_{\Lambda\uparrow\mathbf{R}^2}\frac{1}{|\Lambda|}\,\mathrm{Tr}_\Lambda(f_{T,\mu}(H_L))=-\imath qn\,\frac{\mathscr{E}}{\mathscr{B}}.$$

This is nothing but the classical formula (2), written in complex notation.

Therefore we see that the classical Hall formula still holds in quantum mechanics for the free fermion gas at all temperatures. There is no way to see any trace of quantization of the Hall conductance, neither is there any kind of plateaux of the Hall conductance!

D. The rôle of localization

The main experimental property we have not yet considered is the vanishing of the direct conductivity when the filling factor corresponds to the plateaux. In the previous argument we ignored all effects leading to a finite nonzero direct conductivity. Among these effects there are several sources of dissipation such as phonon scattering or photoemission. At very low temperature, these sources usually have very limited influence and impurity scattering dominates. This is why we are led to consider nondissipative effects like Anderson localization in disordered systems. As explained in Sec. II B, several kinds of defects can influence the electron motion. These defects are usually distributed in a random way so that the forces they create on the charge carriers are actually represented by a random potential. In many cases one considers these defects as isolated and of small influence, so that a first-order perturbation theory based upon one-electron scattering on one impurity already gives a good account of the observed effects. For 2D systems however, it turns out that the low-density limit for impurities does not give the relevant contribution. We will see in Sec. III E how to define properly a potential representing the effects of a high density of random scatterers. For the moment let us stay at an intuitive level.

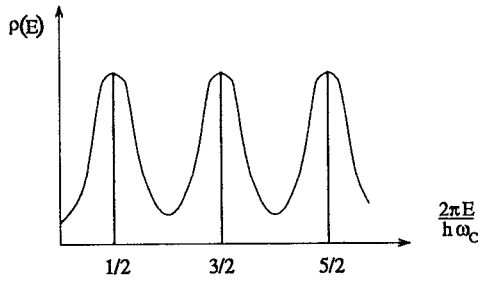

FIG. 3. The density of states for the Landau Hamiltonian with a random potential.

As explained by Anderson,[28] the occurrence of a random potential in a one-particle Hamiltonian may lead to the quantum localization of particles. More precisely, the quantal wave representing these particles reflects on the potential bumps producing interferences. The Bragg condition is necessary for building a constructive interference pattern throughout the crystal. This requires some regularity of the crystal, such as periodicity or quasiperiodicity, in order to allow the wave to propagate. If the potential creating these bumps exhibits some randomness, the probability for the Bragg condition to be satisfied everywhere in the crystal eventually vanishes. Therefore, the wave function will vanish at infinity leading to the trapping of particles in the local minima of the potential. This is Anderson localization.

It has been argued[29] that one-dimensional systems of noninteracting particles in a disordered potential exhibit localization at any strength of disorder. Mathematicians have proved such a claim under relatively mild conditions on the randomness of the potential.[30–34] They have also proved localization in any dimension for strong disorder.[35–40] A finite-size scaling argument, proposed by Thouless,[41] has been used in Ref. 29 to show that two-dimensional systems are also localized at any disorder unless there are spin-orbit couplings.[42] The same argument also shows that in higher dimensions, localization disappears at low disorder. We remark that the notion of localization we are using does not exclude divergence of the localization length at an isolated energy. These results were supplemented by many numerical calculations.[43–46]

On the other hand, the occurrence of a random potential will create new states with energies in the gaps between Landau levels. This can be measured by the *density of states* (DOS). To define it, let $\mathcal{N}(E)$ be the *integrated density of states* (IDS), namely, the number of eigenstates of the Hamiltonian per unit volume below the energy E:

$$\mathcal{N}(E)=\lim_{\Lambda\to\infty}\frac{1}{|\Lambda|}\#\{\text{eigenvalues of } H|_\Lambda\leqslant E\}, \tag{6}$$

where degenerate eigenvalues are counted with their multiplicity. This is a nondecreasing function of E. Therefore its derivative $\rho(E)=d\mathcal{N}(E)/dE$ is well defined as a Stieljes–Lebesgue positive measure and is called the DOS. Under mild conditions on the distribution of the random potential, it is possible to show[47,34,33] that the DOS is a smooth function (see Fig. 3). If the potential strength exceeds the energy difference $\hbar\omega_c$ between two consecutive Landau levels, the gaps will be entirely filled. This is actually the physical situation in most samples used up to now. In heterojunctions, for instance, the minimum of the density of states may represent as much as 30% of its maximum,[21] although in modern samples it is usually very small.[22]

Recall that the chemical potential at zero temperature is called the *Fermi level* E_F. Since the charge carriers are spinless fermions, their density at $T=0$ is given by

$$n=\int_{-\infty}^{E_F} dE\ \rho(E)=\mathcal{N}(E_F). \tag{7}$$

The absence of spectral gaps means that the IDS is monotone increasing, so that changing continuously the particle density is equivalent to changing continuously the Fermi level. More generally, if a magnetic field is switched on, changing continuously the filling factor is equivalent to changing continuously the Fermi level. On the other hand, while the Fermi level crosses a region of localized states, the direct conductivity must vanish whereas the Hall conductivity cannot change since localized states do not participate to the current. This last fact is not immediately obvious, but more justification will be given later on. Changing the filling factor therefore will force the Fermi level to change continuously within this region of localized states while the Hall conductance will stay constant. This is the main mechanism leading to the existence of plateaux.

In contrast, if the spectrum had no localized states, the Hall conductance would change while changing the filling factor, as long as the Fermi level would move within the spectrum. Moreover, if a spectral gap occurred between two bands, let us say between energies E_- and E_+, the IDS would equal a constant n_0 on that gap and changing the filling factor from $n_0-\epsilon$ to $n_0+\epsilon$ would cause the Fermi level to jump discontinuously from E_- to E_+ with the value $(E_-+E_+)/2$ whenever $n=n_0$. In this case, once again, no plateaux could be observed. This is why there is no quantized Hall effect in the free fermion theory.

One of the consequences of this argument is that between two different plateaux, there must be an energy for which the localization length diverges.[51,13] Even though there should be no extended states in 2D disordered systems, the localization length need not be constant in energy. As it happens, if the impurities are electrically neutral in the average, the localization length diverges exactly at the Landau levels. Therefore, the influence of impurities decreases near the band centers, and other sources of interactions, like the Coulomb potential between charge carriers, may become dominant. This is actually the basic observation leading to the understanding of the FQHE.

We have argued that the direct conductivity should vanish whenever the Fermi level lies in a region of localized states. One may wonder how it is then possible to have a nonzero Hall conductance. The answer is that the Hall current is actually carried by edge states. This has been seen in numerical calculations[48,49] and there are the first theoretical hints.[50] The experimental situation is not as clear as it looks from theory. This is probably due to the existence of zones in which states are localized surrounded by filamentary regions in which the Hall current can actually flow.[117]

E. The Laughlin argument

The first attempt to explain the integer quantization of the Hall conductance in units of e^2/h was proposed by Laughlin.[7] Laughlin originally chose a cylinder geometry for his argument. His justification for this was that the effect seemed to be universal and should therefore be independent of the choice of the geometry. Here, we present a Laughlin argument in the plane in form of a singular gauge transformation. In Sec. IV H, this presentation and an observation of Avron, Seiler, and Simon[17] will allow us to discuss the links between the Laughlin argument and the Chern-character approach presented in this article.

Let us consider a free gas of noninteracting electrons in the two-dimensional plane subjected to an exterior magnetic field perpendicular to the plane. We choose an origin and then pass an infinitely thin flux tube through it. A radial electromotive force is created by means of a slowly varying flux $\phi(t)$ (see Fig. 4). In polar coordinates, the vector potential with symmetric gauge for the constant field is

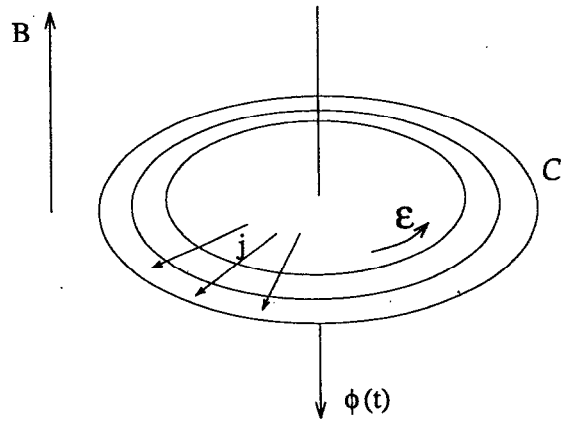

FIG. 4. The chosen geometry for the Laughlin Gedanken experiment.

$$A(r,\theta)=\left(-\frac{B}{2}r\sin\theta-\frac{1}{2\pi r}\sin\theta\,\phi(t),\quad\frac{B}{2}r\cos\theta+\frac{1}{2\pi r}\cos\theta\,\phi(t)\right);$$

the Hamiltonian then reads:

$$H_{\mathscr{B}}=\frac{1}{2m_*}\left\{-\hbar^2\partial_r^2-\hbar^2\frac{1}{r}\partial_r+\left(\frac{\hbar}{\imath}\frac{1}{r}\partial_\theta+\frac{er\mathscr{B}}{2}+\frac{e\phi(t)}{2\pi r}\right)^2\right\}. \tag{8}$$

Recall that the electron charge is $-e$. We now assume that the time evolution of the flux through the tube is so slow that the adiabatic approximation describes the evolution of the states. The eigenstates may then be computed explicitly:

$$\psi_{n,m}(z,\theta;t)=C_{n,m,\phi}e^{-im\theta}\left(\frac{r}{\ell_{\mathscr{B}}}\right)^{m+2\pi e\phi(t)/h}e^{-r^2/4\ell_{\mathscr{B}}^2}L_n^{m+2\pi e\phi(t)/h}\left(\frac{r^2}{\ell_{\mathscr{B}}^2}\right), \tag{9}$$

where $\ell_{\mathscr{B}}=\sqrt{2\hbar/q\mathscr{B}}$ is the magnetic length; $n\in\mathbf{N}$ is the principal quantum number corresponding to the energy $E_n=\hbar\omega_c(n+1/2)$ and $m\in\mathbf{Z}$ is the orbital quantum number so that the angular momentum is $m\hbar$; $C_{n,m,\phi}$ are easily computable normalization constants and the polynomials that appear are the Laguerre polynomials given by:

$$L_n^\alpha(x)=\frac{1}{n!}e^x x^{-\alpha}\partial_x^n(e^{-x}x^{n+\alpha}),\quad n\in\mathbf{N},\quad \alpha\in\mathbf{R}.$$

As the flux $\phi(t)=ht/e\tau$ is varied from $t=0$ to $t=\tau$, one flux quantum is forced through the tube and a state $\psi_{n,m}$ evolves to a state $\psi_{n,m+1}$ up to a phase factor $e^{-i\theta}$. Note that this reflects the fact that the singular gauge transformation introduces an extra angular momentum $-\hbar$ into the system. Now we assume that the filling factor is an integer N. The net effect of the above process is that the state with lowest angular momentum of each Landau level is transported to infinity. Let us fix a large circle $\mathscr{C}$ around the origin of radius R. The current density during the process is then (approximatively) given by $j=1/2\pi R(-Ne/\tau)$; the strength of the electric field on the circle is $\mathscr{E}=-\partial_t(\phi(t)/(2\pi R))=-\hbar/(\tau eR)$; therefore, the Hall conductivity is calculated as $\sigma_H=j/\mathscr{E}=Ne^2/h$.

What did we learn from this argument? Until now, not so much in fact. For had we taken another noninteger filling factor ν and supposed that the electron density were uniform, we would obtain $\sigma_H = \nu e^2/h$; that is the classical result already calculated in Sec. II D.

Using ideas of Prange, Halperin, and Joynt[9,10,51] however, the picture may be completed to furnish a qualitative understanding of the IQHE in the following way: if the sample has impurities, these transform some of the extended states (9) into localized states; moreover, the explicit form of the remaining extended states changes. Now the localized states will not change as a flux quantum is forced through the flux tube, and it can be argued that the remaining extended states of one Landau level still carry the same current, that is one charge e through the circle $\mathscr{C}$. As some of the localized states caused by the impurities are situated below the energy of the Landau levels and others above, we come to the desired qualitative explanation with help of the argument explained in Sec. II D.

There are two ingredients in this Gedanken-experiment which ought to be emphasized. The first is the gauge invariance which produces the periodicity of the Hamiltonian with respect to the varying flux, the period being the flux quantum. The second is that the Hall effect corresponds to a charge transport of one unit for each filled Landau level. The conductance quantization is therefore likely to be connected to the charge quantization.

F. The Chern-Kubo relationship

The geometrical origin of the Hall conductance quantization was revealed by TKN_2 (Ref. 8) and Avron *et al.*[12] TKN_2 considered an electron gas submitted to a uniform magnetic field on a 2D square lattice in the tight-binding approximation. In the Landau gauge the Hamiltonian reads

$$(H\psi)(m_1,m_2) = \psi(m_1-1,m_2) + \psi(m_1+1,m_2) + e^{-2\pi\alpha m_1}\psi(m_1,m_2-1) + e^{2\pi\alpha m_1}\psi(m_1,m_2+1), \tag{10}$$

where $\alpha = \phi/\phi_0$ is the ratio of the magnetic flux ϕ in the unit cell to the flux quantum $\phi_0 = h/e$. Using the translation invariance along the y axis, the solution of the eigenvalue equation can be written in the form

$$\psi(m_1,m_2) = e^{-ik_2 m_2}\varphi(m_1),$$

with

$$\varphi(m_1-1) + \varphi(m_1+1) + 2\cos(k_2 - 2\pi\alpha m_1)\varphi(m_1) = E(k_2)\varphi(m_1). \tag{11}$$

This is Harper's equation.[52] It can also be obtained by adding a weak periodic potential to the Landau Hamiltonian (5) and then projecting onto one Landau level. Now TKN_2 assumed that the ratio α is a rational number of the form p/q where p and q are relatively prime integers. By translation invariance of period q, the solution of Eq. (11) can also be written in Bloch's form

$$\varphi(m_1) = e^{-ik_1 m_1}\xi(m_1), \quad \xi(m_1+q) = \xi(m_1),$$

so that Harper's equation now becomes

$$(H(k_1,k_2)\xi)(n) := e^{ik_1}\xi(n-1) + e^{-ik_1}\xi(n+1) + 2\cos\left(k_2 - 2\pi\frac{p}{q}n\right)\xi(n) = E(k_1,k_2)\xi(n).$$

This is the secular equation for the $q\times q$ Hermitian matrix $H(k_1,k_2)$. The energy spectrum is given by energy bands corresponding to the eigenvalues $\epsilon_l(k_1,k_2)$, $1\leqslant l\leqslant q$. Let $P_l(k_1,k_2)$ be the corresponding eigenprojection. We will assume for simplicity that all bands are well separated: for the Harper equation, all gaps but the central one are open.[53–55]

It is quite clear that $H(k_1,k_2)$ is a trigonometric polynomial in (k_1,k_2), implying that the eigenprojections and the eigenvalues are smooth periodic functions of (k_1,k_2). Moreover, introducing the two unitary $q\times q$ matrices defined by

$$(u_1\xi)(n)=\xi(n-1),\quad (u_2\xi)(n)=e^{-2\imath\pi\frac{p}{q}n}\xi(n),$$

we derive the covariance relation

$$u_1H(k_1,k_2)u_1^{-1}=H\left(k_1,k_2+2\pi\frac{p}{q}\right),\quad u_2H(k_1,k_2)u_2^{-1}=H\left(k_1-2\pi\frac{p}{q},k_2\right). \tag{12}$$

These relations show that the eigenvalues and the trace of any function of H and its derivatives are actually $2\pi/q$-periodic in (k_1,k_2).

The Hall conductance can now be calculated by Kubo's formula.[56] It gives the current as the velocity–velocity correlation function. The formula may be deduced from classical arguments using the Boltzmann equation,[56] but a deduction closer to quantum mechanics using linear response theory will be presented in Sec. IV. The result derived by TKN_2 is then the following: at zero temperature, if the Fermi level belongs to a gap of the Harper Hamiltonian of Eq. (10), the transverse conductivity is given by the Kubo–Chern relation (see Ref. 12)

$$\sigma_H=\frac{e^2}{h}\,\mathrm{Ch}(P_F),$$

where

$$\mathrm{Ch}(P_F)=\int_{\mathrm{T}^2}\frac{d^2k}{4\pi^2}\left\{2\pi\imath\,\frac{1}{q}\,\mathrm{Tr}(P_F(k_1,k_2)[\partial_1P_F(k_1,k_2),\partial_2P_F(k_1,k_2)])\right\}. \tag{13}$$

Here $\partial_a=\partial/\partial k_a$, $a=1,2$, and P_F is the eigenprojection on energies smaller than the Fermi level. The integrand is a complex 2-form on the torus; its de Rham-cohomology class is called the first *Chern class* or the first *Chern character* of the projection P_F. We shall also refer to the integral over this form as Chern character, because the integral will retain meaning in the noncommutative context; the integral will also be called Chern number.

For further informations on Chern classes and characters we refer to Refs. 57 and 58. The Chern character of a projection is a topological invariant. To see this, we remark that the data of the family $k=(k_1,k_2)\in\mathrm{T}^2\mapsto P_F(k)$ of projections defines a complex fiber bundle $\mathscr{P}_F$ over the 2-torus in a natural way. More precisely, $\mathscr{P}_F$ is the set of pairs $(k,\xi)\in\mathrm{T}^2\times\mathbf{C}^q$ such that $P_F(k)\xi=\xi$. Since the Fermi level lies in a gap, P_F is smooth and its dimension is constant. Therefore $\mathscr{P}_F$ is a smooth vector bundle over the 2-torus. It can be shown that $\mathrm{Ch}(P_F)$ depends only upon the equivalence class of $\mathscr{P}_F$ provided the equivalence is isomorphism of vector bundles modulo adding trivial bundles. In particular, it is a homotopy invariant quantity. This implies that adding a small perturbation to the Harper Hamiltonian will not change the value of $\mathrm{Ch}(P_F)$, at least as long as the Fermi level does not cross the spectrum while turning on the perturbation! This is the argument that was needed to explain the robustness of the Hall conductance.

In addition to being robust, the Chern number $\mathrm{Ch}(P_F)$ is actually an integer. To see this explicitly, we will show (see Sec. IV E) that the Chern character Ch is additive with respect to the direct sum of two orthogonal projections P and Q, that is $\mathrm{Ch}(P\oplus Q)=\mathrm{Ch}(P)+\mathrm{Ch}(Q)$. Even

though this is not obvious from Eq. (13), all cross-terms in the left hand side vanish. It is thus sufficient to show that if P is a one dimensional smooth projection over T^2, then its Chern character is an integer. This can be shown as follows.

Let $\xi(k)$ be a unit vector in the vector space $\mathbf{C}^q$ such that $P(k)\xi(k)=\xi(k)$, $\forall k$. It is easy to show that

$$\mathrm{Ch}(P)=\int_{T^2}\frac{d^2k}{2\imath\pi}\,\mathrm{Tr}(P(k)[\partial_1 P(k),\partial_2 P(k)])=\frac{1}{\pi}\int_0^{2\pi}dk_1\int_0^{2\pi}dk_2\,\Im m\langle\partial_1\xi(k)|\partial_2\xi(k)\rangle, \tag{14}$$

where $\Im m$ is the imaginary part. Using Stokes formula, this double integral can be written as

$$\mathrm{Ch}(P)=\int_0^{2\pi}\frac{dk_1}{2\imath\pi}\langle\xi(k)|\partial_1\xi(k)\rangle\Big|_{k_2=0}^{k_2=2\pi}-\int_0^{2\pi}\frac{dk_2}{2\imath\pi}\,\langle\xi(k)|\partial_2\xi(k)\rangle\Big|_{k_1=0}^{k_1=2\pi}. \tag{15}$$

Even though P is doubly periodic in k, ξ need not be periodic. The obstruction to the periodicity of ξ is precisely the nonvanishing of the Chern character. However, we can find two real functions θ_1 and θ_2 such that

$$\xi(k_1,2\pi)=e^{\imath\theta_1(k_1)}\xi(k_1,0),\quad \xi(2\pi,k_2)=e^{\imath\theta_2(k_2)}\xi(0,k_2), \tag{16}$$

provided $0\leqslant k_1\leqslant 2\pi$ and $0\leqslant k_2\leqslant 2\pi$, respectively. The reason is that after one period the new vector ξ defines the same subspace as the old one, so they are proportional. Since $\xi(k)$ is a unit vector for any k, the proportionality factor must be a phase. Writing $\xi(2\pi,2\pi)$ in two ways we get $\theta_1(2\pi)-\theta_1(0)=\theta_2(2\pi)-\theta_2(0)\ (\mathrm{mod}\ 2\pi)$. Replacing in Eq. (15) leads to

$$\mathrm{Ch}(P)=\frac{(\theta_1(2\pi)-\theta_1(0))-(\theta_2(2\pi)-\theta_2(0))}{2\pi}\in\mathbf{Z}.$$

Thanks to this argument, we see that the Chern character of a smooth projection valued function over the 2-torus is an integer.

We remark, however, that the expression used in Eq. (14) is not exactly the same as the one used in Eq. (13): in the latter, we have divided by q. However, the covariance relation Eq. (12) shows that the integrand in Eq. (13) is $2\pi/q$-periodic in k so that it is sufficient to integrate over the square $[0,2\pi/q]^{\times 2}$ and multiply the result by q^2. On the other hand, the periodicity condition (16) has to be modified to take the covariance into account. Now the argument goes along the same line and gives rise to an integer as well.

This series of arguments shows that indeed the Hall conductance (in units of e^2/h) is a very robust integer, as long as the Fermi level remains in a spectral gap. However, the previous argument does not work for any kind of perturbation. It is valid provided the perturbation of the Hamiltonian is given by a matrix valued function $\delta H(k)$ satisfying the same covariance condition (12). Otherwise, the previous construction does not tell us anything about the homotopy invariance of the Hall conductance. This is exactly the limitation of the TKN_2 result. More precisely, no disorder is allowed here and, in addition, only a rational magnetic field can be treated in this way, a quite unphysical constraint. For indeed, in most experiments the variation of the filling factor is obtained through changing continuously the magnetic field. Even though rational numbers are dense in the real line, the previous expression (13) depends in an explicit way on the denominator q of the fraction so that we have no way to check whether the integer obtained is robust when the magnetic field is changed.

The drastic condition of having no disorder is also disastrous: the periodicity of the Hamiltonian excludes localized states, and the argument in Sec. II D shows that no plateaux can be observed in this case. Therefore we have not yet reached the goal.

Nevertheless, the main result of TKN_2 is the recognition that the Hall conductance can be interpreted as a Chern character. This is a key fact in understanding why it is quantized and robust as well. Moreover, the topology underlying this Chern character is that of the quasimomentum space, namely, the periodic *Brillouin zone* which is a 2-torus. From this point of view, the Laughlin argument in its original form may be misleading since it gives the impression of emphasizing the topology of the sample in real space, which has nothing to do with the IQHE.

The solution to the TKN_2 limitation is given by the noncommutative geometry of the Brillouin zone where both of the above restrictions can be dropped. We will see that the Kubo–Chern relation still holds, that the Chern character is still a topological invariant and that it is an integer. But we will also discover something new, characteristic of noncommutative geometry, namely, that localization can be expressed in a very clear way in this context and that the conclusion will be valid as long as the Fermi level lies in a region of *localized states*.

III. THE NONCOMMUTATIVE GEOMETRY OF THE IQHE: THE FOUR-TRACES WAY

In this section the strategy and the main steps of the noncommutative approach are given. In particular we will introduce the four different traces that are technically needed to express the complete results of this theory. The first one is the usual trace on matrices or on trace-class operators. The second one, introduced in the Sec. III A below, is the *trace per unit volume* which permits to compute the Hall conductance by the Kubo formula. The third one is the *graded trace* or *supertrace* introduced in Sec. III B. This is the first technical tool proposed by Connes[4] to define the cyclic cohomology and constitutes the first important step in proving quantization of the Hall conductance.[16] The last one is the *Dixmier trace* defined by Dixmier in 1964[59] and of which the importance for quantum differential calculus was emphasized by Connes.[60,5,6] It will be defined in Sec. IV D but we already explain in Sec. III C below how to use it in order to relate localization properties to the validity of the previous results.

A. The noncommutative Brillouin zone

In a D-dimensional perfect crystal, the description of observables is provided by Bloch's theory. Owing to the existence of a translation group symmetry, each observable of interest commutes with this group. It is a standard result in solid state physics that such an observable is a matrix valued periodic function of the quasimomenta k. The matrix indices usually label both the energy bands and the ions in the unit cell of the crystal whenever it is not a Bravais lattice. The period group in the quasimomentum space is the reciprocal lattice.

If the crystal suffers disorder or if a magnetic field is turned on, this description fails. In most situations however, physicists have found ways to overcome this difficulty. For instance, impurities are treated as isolated objects interacting with the electron Fermi sea using perturbation theory. Magnetic fields in 3-dimensional real crystals are usually so small that a quasiclassical approximation gives a very good account of the physical properties.

Still the conceptual problem of dealing with the breaking of translation symmetry remains. It has not been considered seriously until new physical results forced physicists to face it. One important example in the past was Anderson localization due to a random potential.[28,29,33,32–34] Even though important progress has been made, the solution is still in a rather rough shape and numerical results are still the only justification of many intuitive ideas in this field. When arriving at the QHE, the rôle of localization became so important that there was no way to avoid it.

The main difficulty in such cases is that translates of the one-electron Hamiltonian no longer commute with the Hamiltonian itself. Nevertheless, the crystal under study is still macroscopically homogeneous in space so that its electronic properties are translation invariant. Thus choosing one among the translates of the Hamiltonian is completely arbitrary. In other words, the one-electron Hamiltonian is only known up to translation. Our first proposal to deal with this arbitrariness is to

consider all the translates at once. More precisely, our basic object is the observable algebra $\mathscr{A}$ generated by the family of all translates of the one-electron Hamiltonian in the space G, where G is either $\mathbf{R}^D$ or $\mathbf{Z}^D$.

It is remarkable that the algebra $\mathscr{A}$ exists in the periodic case as well and coincides then with the algebra of matrix-valued, periodic, continuous functions of the quasimomenta.[61] In this case, due to the periodicity with respect to the reciprocal lattice, it is sufficient to consider quasimomenta in the first Brillouin zone $\mathbf{B}$. Then $\mathbf{B}$ is topologically a D-dimensional torus. Therefore the matrix elements of a typical observable are continuous functions over $\mathbf{B}$.

Proceeding by analogy, we will consider the algebra $\mathscr{A}$ in the nonperiodic case as the noncommutative analog of the set of continuous functions over a virtual object called the "noncommutative Brillouin zone."[14,61] This is, however, ineffective as long as we have not defined rules for calculus. In the periodic case physical formulas require two kinds of calculus operations, namely, integrating over the Brillouin zone and differentiating with respect to quasimomenta.

If A is an observable in a periodic crystal, its average over quasimomenta is given by

$$\langle A\rangle = \int_{\mathbf{B}} \frac{d^D k}{|\mathbf{B}|} \operatorname{Tr}(A(k)).$$

Here $|\mathbf{B}|$ represents the volume of the Brillouin zone in momentum space. It turns out that such an average coincides exactly with the "trace per unit volume," namely, we have

$$\langle A\rangle = \lim_{\Lambda\to\infty} \frac{1}{|\Lambda|} \operatorname{Tr}_\Lambda(A_\Lambda) = \mathscr{T}(A), \tag{17}$$

where Λ is a square centered at the origin, $|\Lambda|$ is its volume, $\operatorname{Tr}_\Lambda$ and A_Λ are the restrictions of the trace and the observable A respectively to Λ.

In a nonperiodic, but homogeneous crystal it is still possible to define the trace per unit volume $\mathscr{T}(A)$ of an observable A. Formula (17) however becomes ambiguous in this case, because the limit need not exist. We will show in Sec. III F how to define $\mathscr{T}$ in general. It will satisfy the following properties:

(i) $\mathscr{T}$ is linear, like the integral;
(ii) $\mathscr{T}$ is positive, namely, if A is a positive observable, $\mathscr{T}(A)\geqslant 0$;
(iii) $\mathscr{T}$ is a trace, namely, even though observables may not commute with each other, we still have $\mathscr{T}(AB)=\mathscr{T}(BA)$.

Differentiating with respect to quasimomenta can be defined along the same line. In the periodic case, we remark that the derivative $\partial_j A=\partial A/\partial k_j$ is also given by

$$\partial_j A = -i[X_j, A], \tag{18}$$

where $\mathbf{X}=(X_1,...,X_D)$ is the position operator. We will use the formula (18) as a definition of the derivative in the nonperiodic case. Owing to the properties of the the commutator it satisfies:

(i) ∂_j is linear;
(ii) ∂_j obeys the Leibniz rule $\partial_j(AB)=(\partial_j A)B+A(\partial_j B)$;
(iii) ∂_j commutes with the adjoint operation, namely, $\partial_j(A^*)=(\partial_j A)^*$.

Such a linear map on the observable algebra is called a " *-derivation." In our case it is moreover possible to exponentiate it, for, if $\boldsymbol{\theta}=(\theta_1,...,\theta_D)$, we find

$$e^{\boldsymbol{\theta}\cdot\nabla}(A)=e^{-i\boldsymbol{\theta}\cdot\mathbf{X}}Ae^{i\boldsymbol{\theta}\cdot\mathbf{X}},$$

where $\boldsymbol{\theta}\cdot\mathbf{X}=(\theta_1 X_1+\cdots+\theta_D X_D)$ and $\boldsymbol{\theta}\cdot\nabla=(\theta_1\partial_1+\cdots+\theta_D\partial_D)$. Equipped with the trace per unit volume and with these rules for differentiations, $\mathcal{A}$ becomes a "noncommutative manifold," namely, the *noncommutative Brillouin zone.*

B. Hall conductance and noncommutative Chern character

Using the dictionary created in the previous section between the periodic crystals and the non-periodic ones and in view of Eq. (13), one is led to the following formula for the Hall conductance at zero temperature

$$\sigma_H=\frac{e^2}{h}\,\mathrm{Ch}(P_F),\quad \mathrm{Ch}(P_F)=\frac{1}{2\imath\pi}\,\mathcal{T}(P_F[\partial_1 P_F,\partial_2 P_F]),\tag{19}$$

where P_F is the eigenprojection of the Hamiltonian on energies smaller than or equal to the Fermi level E_F. We will justify this formula in Sec. IV B below within the framework of the relaxation time approximation. However this formula is only valid if

(i) the electron gas is two-dimensional (so $D=2$);
(ii) the temperature is zero;
(iii) the thermodynamic limit is reached;
(iv) the electric field is vanishingly small;
(v) the collision time is infinite;
(vi) electron–electron interactions are ignored.

We will discuss in Sec. IV C the various corrections to that formula whenever these conditions are not strictly satisfied.

The main result of this paper is that *the noncommutative Chern character* $\mathrm{Ch}(P_F)$ *of the Fermi projection is an integer provided the Fermi level belongs to a gap of extended states* (see Theorem 1). Moreover, we will show that *it is a continuous function of the Fermi level as long as this latter lies in a region of localized states* (see Theorem 1). This last result implies that changing the filling factor creates plateaux corresponding to having the Fermi level in an interval of localized states (see Proposition 14). In addition, whenever the Hall conductance jumps from one integer to another one, the localization length must diverge somewhere in between [see Corollary (3)].

In order to get this result we will follow a strategy introduced by Connes in this context,[4] but originally due to Atiyah.[62] Before describing it, we need some notations and definitions. We build a new Hilbert space which is made of two copies $\mathcal{H}_+$ and $\mathcal{H}_-$ of the physical Hilbert space $\mathcal{H}$ of quantum states describing the electron. This is like working with Pauli spinors. In this doubled Hilbert space $\hat{\mathcal{H}}=\mathcal{H}_+\oplus\mathcal{H}_-$, the grading operator $\hat{G}$ and the "Hilbert transform" F are defined as follows:

$$\hat{G}=\begin{pmatrix}+1 & 0\\ 0 & -1\end{pmatrix},\quad F=\begin{pmatrix}0 & \dfrac{X}{|X|}\\ \dfrac{\bar{X}}{|X|} & 0\end{pmatrix},\tag{20}$$

where $X=X_1+\imath X_2$ (here the dimension is $D=2$). It is clear that F is self-adjoint and satisfies $F^2=\mathbf{1}$. An operator T on $\hat{\mathcal{H}}$ will said to be of degree 0 if it commutes with $\hat{G}$ and of degree 1 if it anticommutes with $\hat{G}$. Then, every operator on $\hat{\mathcal{H}}$ can be written in a unique way as the sum $T=T_0+T_1$ where T_j has degree j. The graded commutator (or supercommutator) of two operators and the graded differential dT are defined by

$$[T, T']_S = TT' - (-)^{deg(T)deg(T')}T'T, \quad dT = [F, T]_S.$$

Then, $d^2T = 0$. The graded trace Tr_S (or supertrace) is defined by

$$\mathrm{Tr}_S(T) = \tfrac{1}{2}\,\mathrm{Tr}_{\hat{\mathscr{H}}}(\hat{G}F[F,T]_S) = \mathrm{Tr}_{\mathscr{H}}(T_{++} - uT_{--}\bar{u}), \tag{21}$$

where $u = X/|X|$ and T_{++} and T_{--} are the diagonal components of T with respect to the decomposition of $\hat{\mathscr{H}}$. It is a linear map on the algebra of operators such that $\mathrm{Tr}_S(TT') = \mathrm{Tr}_S(T'T)$. Moreover, operators of degree 1 have zero trace. However, this trace is not positive. Observables in $\mathscr{A}$ will become operators of degree 0, namely, $A \in \mathscr{A}$ will be represented by $\hat{A} = A_+ \oplus A_-$ where $A_\pm$ acts as A on each of the components $\mathscr{H}_\pm = \mathscr{H}$.

With this formalism, Connes[4] gave a formula which was extended to the present situation in Ref. 16 to the following one (see Sec. IV F):

$$\mathrm{Ch}(P_F) = \int_\Omega d\mathbf{P}(\omega)\mathrm{Tr}_S(\hat{P}_F d\hat{P}_F d\hat{P}_F), \tag{22}$$

where the space Ω represents the disorder configurations while the integral over $d\mathbf{P}$ is the average over the disorder. We will show that the integrand does not depend upon the disorder configuration. Moreover, we will show that the right hand side of (22) is **P**-almost surely equal to the index of a Fredholm operator, namely,

$$\mathrm{Ch}(P_F) = \mathrm{Ind}(P_F F_{+-} \restriction_{P_F(\mathscr{H})}). \tag{23}$$

The Chern character is therefore an integer, at least whenever it is well-defined.

One may wonder what is the physical meaning of this integer. Actually, an answer was provided quite recently by Avron *et al.*[17] Considering the definition of the graded trace in formula (21), they interpreted the index found for the Chern character as the relative index of the projections P_F and $uP_F\bar{u}$. This is to say even though these projections are infinite-dimensional, their difference has finite rank giving rise to a new index called the "relative index." Then they argue that u represents the action of the singular Laughlin gauge (see Sec. II E) and that this difference can be interpreted as a charge transported to infinity as in the original Laughlin argument.[7]

C. Localization and plateaux of the Hall conductance

In Sec. III B, we have explained the topological aspect of the Hall conductance quantization. However, these formulas only hold if both sides are well defined. For instance, it is not clear whether the Fermi projection is differentiable. If it is not, what is the meaning of formula (19)? This is the aspect that we want to discuss now.

First of all, using the Schwarz inequality, a sufficient condition for formula (19) to hold is that the Fermi projection satisfy $\mathscr{T}(|\nabla P_F|^2) < \infty$ if we set $\nabla = (\partial_1, \partial_2)$. It is important to remark that this condition is much weaker than demanding P_F to be differentiable. This is rather a noncommutative analog of a Sobolev norm (or the square of it). By definition, this expression can be written as

$$\mathscr{T}(|\nabla P_F|^2) = \int_\Omega d\mathbf{P}(\omega) \int_{x \in G} d^2x |\langle 0|P_F|\mathbf{x}\rangle|^2 |\mathbf{x}|^2. \tag{24}$$

Note that Eq. (24) can be defined in any dimension D. From this expression, we see that the finiteness of the Sobolev norm is related to the finiteness of some localization length. Indeed, we will show that only the part of the energy spectrum near the Fermi level gives a contribution to formula (24) so that P_F can be replaced by the projection P_Δ onto energies in some interval Δ

close to E_F. On the other hand, the matrix element $\langle 0|P_\Delta|\mathbf{x}\rangle$ gives a measure of how far states with energies in Δ are localized. Thus, the Sobolev norm is some measurement of the localization length. We will develop this fact in Sec. V B below. One way of defining a localization length consists in measuring how far a wave packet goes in time. This gives

$$l^2(\Delta)=\lim_{T\to\infty}\frac{1}{T}\int_{t=0}^{T}dt\int_\Omega d\mathbf{P}(\omega)\langle 0||(\mathbf{X}(t)-\mathbf{X})|^2|0\rangle,$$

where the time evolution is governed by HP_Δ and where $|A|^2=\Sigma_i A_i^* A_i$ if $A=(A_1,...,A_D)$. In this expression, $\mathbf{X}$ is the position operator and $\mathbf{X}(t)$ is the evolution of $\mathbf{X}$ under the one-electron Hamiltonian after time t. Let now $\mathcal{N}$ be the density of states (DOS) already introduced in Sec. II D. It can be defined equivalently as the positive measure $\mathcal{N}$ on the real line such that for any interval Δ

$$\mathcal{T}(P_\Delta)=\int_\Delta d\mathcal{N}(E).$$

That these two definitions coincide is guaranteed by a theorem of Shubin.[61,63] We will show that if $l^2(\Delta)$ is finite, there is a positive $\mathcal{N}$-integrable function $l(E)$ over Δ such that

$$l^2(\Delta')=\int_{\Delta'}d\mathcal{N}(E)l(E)^2,$$

for any Borel subset Δ' of Δ (see Theorem 13). The number $l(E)$ will be called the "localization length" at energy E. Moreover, we also find

$$\mathcal{T}(|\nabla P_{\Delta'}|^2)\leqslant\int_{\Delta'}d\mathcal{N}(E)l(E)^2,$$

such that the mapping $E_F\in\Delta\mapsto P_F$ is continuous with respect to the Sobolev norm (Theorem 14).

The previous argument shows that, whenever the localization length at the Fermi level is finite, the Chern character of the Fermi projection is well defined. Now let us consider the right hand side of formula (22). In a recent work Connes[60,5] proposed to use another trace which was introduced in the sixties by Dixmier.[59] This Dixmier trace $\mathrm{Tr}_{\mathrm{Dix}}$ will be defined in Sec. IV D. Let us simply say that it is a positive trace on the set of compact operators, such that any trace-class operator is annihilated by it whereas if T is a positive operator with finite Dixmier trace, then $T^{1+\epsilon}$ is trace class for any $\epsilon>0$. In Ref. 60, Connes proved a formula relating the so-called Wodzicky residue of a pseudodifferential operator to its Dixmier trace. Adapting this formula to our situation leads to the following remarkable result (see Theorem 9):

$$\mathcal{T}(|\nabla P_F|^2)=\frac{1}{\pi}\,\mathrm{Tr}_{\mathrm{Dix}}(|d\hat{P}_F|^2).$$

Note that this formula only holds for electrons in two dimensions. In higher dimensions, things are more involved. Therefore, we see that as soon as the localization length is finite in a neighborhood of the Fermi level, the operator $d\hat{P}_F$ is square summable with respect to the Dixmier trace implying that its third power is trace-class. In view of the formulas (21) and (22), we see that this is exactly the condition for the existence of the right-hand side of Eq. (22). Moreover, the continuity of P_F in E_F with respect to the Sobolev norm implies that the index in formula (23) is

constant as long as the Fermi level stays in a region in which the localization length is finite. Therefore the Chern character is an integer and the Hall conductance is quantized and exhibits plateaux.

In this way, the mathematical frame developed here gives rise to a complete mathematical description of the IQHE, within the approximations that have been described previously.

D. Summary of the main results

Let us summarize our mathematical results in this section.

Theorem 1: *Let $H=H^*$ be a Hamiltonian affiliated to the C^*-algebra $\mathcal{A}=C^*(\Omega\times G,\mathcal{B})$ defined in Sec. III F. Let* **P** *be a G-invariant and ergodic probability measure on Ω. Then we have the following results:*

(i) **(Kubo–Chern formula)** *In the limit where*
(a) the volume of the sample is infinite,
(b) the relaxation time is infinite, and
(c) the temperature is zero,

the Hall conductance of an electron gas without interaction and described by the one-particle Hamiltonian H is given by the formula

$$\sigma_H=\frac{e^2}{h}\,\mathbf{Ch}(P_F)=\frac{e^2}{h}\,2\imath\pi\mathcal{T}(P_F[\partial_1 P_F,\partial_2 P_F]),$$

where P_F is the eigenprojection of H on energies smaller than or equal to the Fermi level E_F and $\mathcal{T}$ is the trace per unit volume associated to **P**.

(ii) **(Quantization of the Hall conductance)** *If in addition P_F belongs to the Sobolev space $\mathcal{S}$ associated to $\mathcal{A}$, then $\mathbf{Ch}(P_F)$ is an integer which represents the charge transported at infinity by a Laughlin adiabatic switching of a flux quantum.*

(iii) **(Localization regime)** *Under the same conditions as the ones in* (ii), *the direct conductivity vanishes.*

(iv) **(Existence of Plateaux)** *Moreover, if Δ is an interval on which the localization length $l^2(\Delta)$ defined in Sec. III C is finite, then as long as the Fermi energy stays in Δ and is a continuity point of the density of states of H, $\mathbf{Ch}(P_F)$ is constant and P_F belongs to the Sobolev space $\mathcal{S}$.*

The following corollary is an immediate consequence. It was proved in Refs. 51 and 13.

Corollary 1: If the Hall conductance σ_H jumps from one integer to another in between the values ν_1 and ν_2 of the filling factor, there is an energy between the corresponding values of the Fermi levels at which the localization length diverges.

Strictly speaking, this theorem has been completely proved only for the case of a discrete lattice (tight-binding representation). Most of it is valid for the continuum, but parts of the proofs require extra technical tools so that the proof of this theorem is not complete in this paper. We postpone the complete proof of it for the continuum case to a future work.[64]

As a side result, we emphasize that we have developed a noncommutative framework valid to justify the transport theory for aperiodic media (see Sec. IV A below). It allows us to give rigorous estimates on the error terms whenever the conditions of the previous theorem are not strictly satisfied (see Sec IV C below). We will not give the mathematical proofs that these errors are

rigorous bounds here even though they actually are. This will also be the main topic of a future work. But we have estimated them and we show that they are compatible with the accuracy observed in the experiments.

Another result which is actually completely developed here due to its importance in the quantum Hall effect, concerns the definition and the properties of the localization length. We give a noncommutative expression for it and show how it is related to the existence of plateaux. The main results in this direction are the Theorems 13 and 14 in Sec. V B. We also show that the localization length is indeed finite in models like the Anderson model for disordered systems for which proofs of exponential localizations are available.

The proof of Theorem I will be divided up into a number of partial results; there will be no explicit paragraph '**Proof of Theorem 1,**' let us therefore outline the main steps. In Secs. III E and III F, we give a precise mathematical description of a homogeneous Schrödinger operator and its hull and we construct the observable algebra. In Sec. IV the Kubo formula for the Hall conductance is derived and we present the (noncommutative) geometrical argument for the integer quantization of the Hall conductance (compare Theorem 11). That the proven index theorem is precisely valid under the weak localization condition $P_F \in \mathcal{S}$ results from Theorem 9. Point (iv) is proved in Sec. V.

E. Homogeneous Schrödinger's operators

Most of the results of the next two sections have already been proved in Ref. 61, Sec. 2. We will only give the main steps that are necessary in this paper for the purpose of the IQHE.

In earlier works,[14,61] one of us has introduced the notion of a homogeneous Hamiltonian. The purpose of this notion is to describe materials which are translation invariant at a macroscopic scale but not necessarily at a microscopic one. In particular, it is well suited for the description of aperiodic materials. In such a medium, there is no natural choice of an origin in space. If H is a Hamiltonian describing one particle in this medium, we can replace it by any of its translates $H_a = U(a)HU(a)^{-1}, a \in \mathbf{R}^D$; the physics will be the same. This choice is entirely arbitrary, so that the smallest possible set of observables must contain at least the full family $\{H_a; a \in \mathbf{R}^D\}$; this family will be completed with respect to a suitable topology. We remark that H need not be a bounded operator, so that calculations are made easier if we consider its resolvent instead. Let us define precisely what we mean by "homogeneity" of the medium described by H.

Definition 1: Let $\mathcal{H}$ be a separable Hilbert space. Let G be a locally compact group (for instance $\mathbf{R}^D$ or $\mathbf{Z}^D$). Let U be a unitary projective representation of G, namely, for each $a \in G$, there is a unitary operator $U(a)$ acting on $\mathcal{H}$ such that the family $U = \{U(a); a \in G\}$ satisfies the following properties:

(i) $U(a)U(b) = U(a+b)e^{i\phi(a,b)}$ $\forall a,b \in G$, where $\phi(a,b)$ is some phase factor.
(ii) For each $\psi \in H$, the map $a \in G \mapsto U(a)\psi \in \mathcal{H}$ is continuous.

Then a self-adjoint operator H on $\mathcal{H}$ is homogeneous with respect to G if the family $\{R_a(z) = U(a)(z\mathbf{1} - H)^{-1}U(a)^{-1}; a \in G\}$ admits a compact strong closure.

Remark. A sequence $(A_n)_{n>0}$ of bounded operators on $\mathcal{H}$ converges strongly to the bounded operator A if for every $\psi \in \mathcal{H}$, the sequence $\{A_n\psi\}$ of vectors in $\mathcal{H}$ converges in norm to $A\psi$. The set considered in the definition has therefore a strong compact closure if, for given $\varepsilon > 0$ and for a finite set $\psi_1, \ldots, \psi_N$ of vectors in $\mathcal{H}$, there is a finite set $a_1, \ldots, a_m$ in G such that for every $a \in G$ and every $1 \leq j \leq N$, there is $1 \leq i \leq m$ such that $\|(R_a(z) - R_{a_i}(z))\psi_j\| \leq \varepsilon$. In other words, the full family of translates of $R(z)$ is well approximated on vectors by a finite number of these translates; this finite number then repeats itself infinitely many times up to infinity.

The virtue of this definition comes from the construction of the "hull." Let z belong to the resolvent set $\rho(H)$ of H and let H be homogeneous. We denote by $\Omega(z)$ the strong closure of the

family $\{R_a(z)=U(a)(z\mathbf{1}-H)^{-1}U(a)^{-1};a\in G\}$; it is therefore a compact space. Moreover, it is endowed with an action of the group G by means of the representation U. This action defines a group of homeomorphisms of $\Omega(z)$. Thanks to the resolvent equation, it is quite easy to prove that if z' is another point in the resolvent set $\rho(H)$, the spaces $\Omega(z)$ and $\Omega(z')$ are homeomorphic.[61] Identifying them gives rise to an abstract compact space Ω endowed with an action of G by a group of homeomorphisms. If $\omega\in\Omega$ and $a\in G$, we will denote by $T^a\omega$ the result of the action of a on ω, and by $R_\omega(z)$ the representative of ω in $\Omega(z)$. Then we have

$$U(a)R_\omega(z)U(a)^*=R_{T^a\omega}(z),$$

$$R_\omega(z')-R_\omega(z)=(z-z')R_\omega(z')R_\omega(z). \tag{25}$$

In addition, $z\mapsto R_\omega(z)$ is norm-holomorphic on $\rho(H)$ for every $\omega\in\Omega$, and the application $\omega\mapsto R_\omega(z)$ is strongly continuous.

Definition 2: Let H be an operator, homogeneous with respect to the representation U of the locally compact group G on the Hilbert space $\mathscr{H}$. Then the hull of H is the dynamical system (Ω,G,T) where Ω is the compact space given by the strong closure of the family $\{R_a(z)=U(a)(z\mathbf{1}-H)^{-1}U(a)^{-1};a\in G\}$, and G acts on Ω through T.

In general, Eq. (25) is not sufficient to ensure that $R_\omega(z)$ is the resolvent of some self-adjoint operator H_ω, for indeed, one may have $R_\omega(z)=0$ if no additional assumption is demanded. A sufficient condition is that H be given by H_0+V where: (i) H_0 is self-adjoint and G-invariant, (ii) V is relatively bounded with respect to H_0, i.e., $\|(z-H_0)^{-1}V\|<\infty$, (iii) $\lim_{|z|\to\infty}\|(z-H_0)^{-1}V\|=0$. Then, $R_\omega(z)=\{\mathbf{1}-(z\mathbf{1}-H_0)^{-1}V_\omega\}^{-1}(z\mathbf{1}-H_0)^{-1}$ where $(z\mathbf{1}-H_0)^{-1}V_\omega$ is defined by the strong limit of $(z\mathbf{1}-H_0)^{-1}V_{a_i}$, which obviously exists. So $R_\omega(z)$ is the resolvent of H_0+V_ω.

If H is a Schrödinger operator, the situation becomes simpler. Let us consider the case of a particle in $\mathbf{R}^2$ with mass m and charge q, submitted to a bounded potential V and a uniform magnetic field $\mathscr{B}$ with vector potential $\mathbf{A}$. We will describe the vector potential in the symmetric gauge, namely $\mathbf{A}=(-\mathscr{B}x_2/2,\mathscr{B}x_1/2)$. The Schrödinger operator is given by

$$H=\frac{1}{2m}\sum_{j=1,2}(P_j-qA_j)^2+V=H_L+V. \tag{26}$$

The unperturbed part H_L is translation invariant, provided one uses magnetic translations[27] defined by [if $\mathbf{a}\in\mathbf{R}^2$, $\psi\in L^2(\mathbf{R}^2)$]

$$U(\mathbf{a})\psi(\mathbf{x})=\exp\left\{\frac{-iq\mathscr{B}}{2\hbar}\mathbf{x}\wedge\mathbf{a}\right\}\psi(\mathbf{x}-\mathbf{a}), \tag{27}$$

where $\mathbf{x}\wedge\mathbf{a}=x_1a_2-x_2a_1$. It is easy to check that the operators $U(a)$ form a projective unitary representation of the translation group $\mathbf{R}^2$. The main result in this case is given by

Theorem 2: *Let H be given by Eq. (26) with V a measurable essentially bounded function on $\mathbf{R}^2$. Then*

(i) H is homogeneous with respect to the magnetic translations (27);
(ii) the hull ω of H is homeomorphic to the hull of V, namely the weak closure of the family $\{V(.-\mathbf{a});\mathbf{a}\in\mathbf{R}^2\}$ in $L^\infty_{\mathbf{R}}(\mathbf{R}^2)$;
(iii) there is a Borelian function v on Ω such that, if we denote by V_ω the bounded function representing the point $\omega\in\Omega$, then $V_\omega(\mathbf{x})=v(T^{-\mathbf{x}}\omega)$ for almost every $\mathbf{x}\in\mathbf{R}^2$ and all $\omega\in\Omega$. If in addition V is uniformly continuous and bounded, then v is continuous.

The proof of this theorem can be found in Ref. 6, Sec. II D.

In many cases, it is actually better to work in the tight-binding representation. The reason is that only electrons with energies near the Fermi level contribute to the current. One usually defines an effective Hamiltonian by reducing the Schrödinger operator to the interval of energies of interest.[14] The Hamiltonian can then be described as a matrix $H(x,x')$ indexed by sites in the lattice $\mathbf{Z}^2$ acting on elements of the Hilbert space $\ell^2(\mathbf{Z}^2)$ as follows

$$H\psi(\mathbf{x})=\sum_{\mathbf{x}'\in\mathbf{Z}^2} H(\mathbf{x},\mathbf{x}')\exp\left\{\imath\pi\,\frac{\phi}{\phi_0}\,\mathbf{x}\wedge\mathbf{x}'\right\}\psi(\mathbf{x}'),\quad \psi\in\ell^2(\mathbf{Z}^2), \tag{28}$$

where ϕ is the flux in the unit cell, whereas $\phi_0=h/q$ is the flux quantum. In most examples, one can find a sequence f such that $\Sigma_{\mathbf{a}\in\mathbf{Z}^2}f(\mathbf{a})<\infty$ and $|H(\mathbf{x},\mathbf{x}')|\leqslant f(\mathbf{x}-\mathbf{x}')$. In particular, H is bounded and there is no longer a need to consider its resolvent. Let now U be the unitary projective representation of the translation group $\mathbf{Z}^2$ given by

$$U(\mathbf{a})\psi(\mathbf{x})=\exp\left\{\imath\pi\,\frac{\phi}{\phi_0}\mathbf{x}\wedge\mathbf{a}\right\}\psi(\mathbf{x}-\mathbf{a}),\quad \psi\in\ell^2(\mathbf{Z}^2). \tag{29}$$

Theorem 3: *Let H be given by Eq. (28). Then H is homogeneous with respect to the projective representation U [Eq. (29)] of the translation group. Moreover, if Ω is the hull of H, there is a continuous function $\hat{h}$ on $\Omega\times\mathbf{Z}^2$, vanishing at infinity, such that $H_\omega(\mathbf{x},\mathbf{x}')=\hat{h}(T^{-\mathbf{x}}\omega,\mathbf{x}'-\mathbf{x})$, for every pair $(\mathbf{x},\mathbf{x}')\in\mathbf{Z}^2$ and $\omega\in\Omega$.*

Again the proof can be found in Ref. 61, Sec. II D.

F. Observables and calculus

In the previous section we have constructed the hull of a Hamiltonian. It is a compact space that represents the degree of aperiodicity of the crystalline forces acting on the charge carriers. For disordered systems, the hull is nothing but the space of disorder configurations. In principle, the algebra of observables should be constructed from the Hamiltonian H by taking all functions of H and its translates. However, we proceed in a different way giving a more explicit construction which may give a bigger algebra in general, but will be easier to use.

Let Ω be a compact topological space equipped with a $\mathbf{R}^2$-action by a group $\{T^a;a\in\mathbf{R}^2\}$ of homeomorphisms. Given a uniform magnetic field $\mathscr{B}$, we can associate to this dynamical system a C^*-algebra $C^*(\Omega\times\mathbf{R}^2,\mathscr{B})$ as follows. We first consider the topological vector space $\mathscr{C}_\kappa(\Omega\times\mathbf{R}^2)$ of continuous functions with compact support on $\Omega\times\mathbf{R}^2$. It is endowed with the following structure of a *-algebra by

$$AB(\omega,\mathbf{x})=\int_{\mathbf{R}^2}d^2y\;A(\omega,\mathbf{y})B(T^{-\mathbf{y}}\omega,\mathbf{x}-\mathbf{y})\exp\left\{\frac{\imath q\mathscr{B}}{2\hbar}\mathbf{x}\wedge\mathbf{y}\right\}, \tag{30}$$

$$A'(\omega,\mathbf{x})=\overline{A(T^{-\mathbf{x}}\omega,-\mathbf{x})},$$

where $A,B\in\mathscr{C}_\kappa(\Omega\times\mathbf{R}^2)$, $\omega\in\Omega$, and $x\in\mathbf{R}^2$. For $\omega\in\Omega$, this *-algebra is represented on $L^2(\mathbf{R}^2)$ by

$$\pi_\omega(A)\psi(\mathbf{x})=\int_{\mathbf{R}^2}d^2y\;A(T^{-\mathbf{x}}\omega,\mathbf{y}-\mathbf{x})\exp\left\{\frac{\imath q\mathscr{B}}{2\hbar}\,\mathbf{y}\wedge\mathbf{x}\right\}\psi(\mathbf{y}),\quad \psi\in L^2(\mathbf{R}^2), \tag{31}$$

namely, π_ω is linear, $\pi_\omega(AB)=\pi_\omega(A)\pi_\omega(B)$ and $\pi_\omega(A)^*=\pi_\omega(A^*)$. In addition, $\pi_\omega(A)$ is a bounded operator for $\|\pi_\omega(A)\|\leqslant\|A\|_{\infty,1}$ where

$$\|A\|_{\infty,1}=\max\left\{\sup_{\omega\in\Omega}\int_{\mathbf{R}^2}d^2y|A(\omega,\mathbf{y})|,\ \sup_{\omega\in\Omega}\int_{\mathbf{R}^2}d^2y|A^*(\omega,\mathbf{y})|\right\}.$$

This defines a norm which satisfies $\|AB\|_{\infty,1}\leq\|A\|_{\infty,1}\|B\|_{\infty,1}$, $\|A\|_{\infty,1}=\|A^*\|_{\infty,1}$. Since A has compact support, its $(\infty,1)$-norm is finite. We will denote by $L^{\infty,1}(\Omega\times\mathbf{R}^2;\mathcal{B})$ the completion of $\mathcal{C}_\kappa(\Omega\times\mathbf{R}^2)$ under this norm. We then remark that these representations are related by the covariance condition

$$U(\mathbf{a})\pi_\omega(A)U(\mathbf{a})^{-1}=\pi_{T^{\mathbf{a}}\omega}(A).$$

Now we set

$$\|A\|=\sup_{\omega\in\Omega}\|\pi_\omega(A)\|,$$

which defines a C^*-norm on $L^{\infty,1}(\Omega\times\mathbf{R}^2;\mathcal{B})$. This permits us to define $C^*(\Omega\times\mathbf{R}^2,\mathcal{B})$ as the completion of $\mathcal{C}_\kappa(\Omega\times\mathbf{R}^2)$ or of $L^{\infty,1}(\Omega\times\mathbf{R}^2;\mathcal{B})$ under this norm. Clearly, the representations π_ω can be continuously extended to this C^*-algebra. We remark that this algebra has no unit.

In particular, for the Hamiltonian (26) we get [Ref. 16, Sec. 2.5]:

Theorem 4: *Let H be given by Eq. (26) where $\mathbf{A}$ is the vector potential of a uniform magnetic field in the symmetric gauge and let V* be in $L^\infty(\mathbf{R}^2)$. *We denote by Ω its hull. Then for each z in the resolvent set of H and for every $\mathbf{x}\in\mathbf{R}^2$, there is an element $R(z;\mathbf{x})\in C^*(\Omega\times\mathbf{R}^2,\mathcal{B})$, such that for each $\omega\in\Omega$, $\pi_\omega(R(z;\mathbf{x}))=(z\mathbf{1}-H_{T^{-\mathbf{x}}\omega})^{-1}$.*

In the discrete case [see Eqs. (28)and (3)], the same construction holds provided we replace $\mathbf{R}^2$ by $\mathbf{Z}^2$ and the integral over $\mathbf{R}^2$ by a sum over $\mathbf{Z}^2$, namely,

Theorem 5: *Let H be given by Eq. (28) where ϕ is the magnetic flux through the unit cell and $\phi_0=h/q$ is the flux quantum. We denote by Ω its hull as in Theorem 3. Then the function $\hat{h}$ appearing in Theorem 3 belongs to $C^*(\Omega\times\mathbf{Z}^2,\mathcal{B})$ and for each $\omega\in\Omega$, $\pi_\omega(\hat{h})=H_\omega$.*

We remark that in the discrete case, the function $\mathbf{1}(\omega,\mathbf{x})=\delta_{\mathbf{x},\mathbf{0}}$ is a unit of $C^*(\Omega\times\mathbf{Z}^2,\mathcal{B})$; this is the main difference between the continuous and the discrete case. In the noncommutative terminology, the discrete case corresponds to a compact noncommutative manifold, whereas the continuous case corresponds to a locally compact, but not compact, noncommutative manifold. In the sequel, many results will hold for both of the constructed C^*-algebras. We therefore introduce the notation $\mathcal{A}$ for both of them. The algebras of functions with compact support which are at the base of the construction are denoted by $\mathcal{A}_0$. Moreover, let G be the physical space; it is either $\mathbf{R}^2$ or $\mathbf{Z}^2$.

If the Hamiltonian H is unbounded, it does not belong to the C^*-algebra $\mathcal{A}$. However, we have seen in Theorem 4 that its resolvent belongs to $\mathcal{A}$. As a consequence, $f(H)$ belongs to $\mathcal{A}$ for every continuous function f on the real line vanishing at infinity. This leads to the following definition.

Definition 3: Let $\mathcal{A}$ be the C^*-algebra $C^*(\Omega\times G,\mathcal{B})$ defined above and let $\mathcal{H}$ be the Hilbert space $L^2(G)$ endowed with the projective representation U defined in Eqs. (27) and (29). We will say that a self-adjoint homogeneous operator H is affiliated to $\mathcal{A}$ whenever its hull can be embedded in the dynamical system (Ω,G,T) and if for each z in its resolvent set, there is an element $R(z)\in\mathcal{A}$ such that

$$(z\mathbf{1}-H_\omega)^{-1}=\pi_\omega(R(z)),$$

for all $\omega\in\Omega$.

By 'embedded' we mean that the hull is a T-invariant subset of Ω.

The two rules of calculus, namely, integration and derivations, are now constructed as follows. First of all, let $\mathbf{P}$ be a G-invariant probability measure on Ω. For most of the results of this paper, the choice of such a measure is irrelevant. We will discuss its physical relevance in Secs. IV G and VI B. A trace on $\mathcal{A}$ is constructed as follows. If $A \in \mathcal{A}_0$ we set

$$\mathcal{T}_{\mathbf{P}}(A) = \int_{\Omega} d\mathbf{P}(\omega) A(\omega, \mathbf{0}).$$

This defines a positive trace (see Sec. III A). It is faithful [namely, $\mathcal{T}_{\mathbf{P}}(AA^*) = 0$ implies $A = 0$] whenever the support of $\mathbf{P}$ is Ω. The trace is normalized if $G = \mathbf{Z}^2$ [namely, $\mathcal{T}_{\mathbf{P}}(\mathbf{1}) = 1$] and is unbounded if $G = \mathbf{R}^2$. It actually coincides with the trace per unit area. More precisely we have[14,61]

Proposition 1: Let A belong to $\mathcal{A}_0$. *Then for* **P***-almost all* ω*'s*

$$\mathcal{T}_{\mathbf{P}}(A) = \lim_{\Lambda \uparrow G} \frac{1}{|\Lambda|} \mathrm{Tr}_{\Lambda}(\pi_{\omega}(A)), \tag{32}$$

where Λ *denotes a sequence of squares in G centered at the origin and covering G and* Tr_{Λ} *is the restriction to* Λ *of the usual trace.*

In the sequel, we will drop the subscript $\mathbf{P}$.

Given a self-adjoint element $H \in \mathcal{A}$, we define its DOS as the positive measure $\mathcal{N}$ on the real line such that for any continuous function f with compact support on $\mathbf{R}$

$$\mathcal{T}(f(H)) = \int_{\mathbf{R}} d\mathcal{N}(E) f(E).$$

In view of (32), this definition agrees with the definition given in Eq. (7).

If $p \geq 1$, we denote by $L^p(\mathcal{A}, \mathcal{T})$ the completion of $\mathcal{A}_0$ under the norm

$$\|A\|_{L^p} = (\mathcal{T}(\{AA^*\}^{p/2}))^{2/p}.$$

In particular, one can check that the space $L^2(\mathcal{A}, \mathcal{T})$ is a Hilbert space (GNS construction) identical to $L^2(\Omega \times G, d\mathbf{P} \otimes dg)$. The map $\phi \in L^2(\mathcal{A}, \mathcal{T}) \mapsto A\phi \in L^2(\mathcal{A}, \mathcal{T})$ for $A \in \mathcal{A}$ defines a representation π_{GNS} of $\mathcal{A}$. The weak closure $L^{\infty}(\mathcal{A}, \mathcal{T}) = \pi_{GNS}(\mathcal{A})''$ is a von Neumann algebra. By construction, the trace $\mathcal{T}$ extends to a trace on this algebra. We remark that if H is a self-adjoint element of $\mathcal{A}$, its eigenprojections are in general elements of the von Neumann algebra $L^{\infty}(\mathcal{A}, \mathcal{T})$.

Let us give another characterization of the von Neumann algebra $L^{\infty}(\mathcal{A}, \mathcal{T})$ which can be found in Ref. 65. Let $\mathcal{W}$ be the set of weakly measurable families $\omega \in \Omega \mapsto A_{\omega}$ of bounded operators on $L^2(G)$ which are covariant

$$U(a) A_{\omega} U(a)^{-1} = A_{T^a \omega}, \quad a \in G, \ \omega \in \Omega$$

and $\mathbf{P}$-essentially bounded. This latter means that the norm of A_{ω} is bounded in ω except possibly on a subset of zero probability with respect to $\mathbf{P}$. We endow $\mathcal{W}$ with the norm

$$\|A\|_{\mathbf{P}} = \mathbf{P}\text{-ess sup}_{\omega} \|A_{\omega}\|_{\mathcal{L}(L^2(G))}, \quad A \in \mathcal{W}.$$

Sum, product, and adjoint of elements of $\mathcal{W}$ are defined pointwise in the obvious way. Then Connes[65] proved that $\mathcal{W}$ is a von Neumann algebra, namely, a C^*-algebra with predual,[66] and that

Theorem 6: *$L^\infty(\mathcal{A},\mathcal{T})$ is canonically isomorphic to the von Neumann algebra $\mathcal{W}$ of* **P***-essentially bounded measurable covariant families of operators on $L^2(G)$.*

Actually, this isomorphism is obvious if we realize that the Hilbert space $L^2(\mathcal{A},\mathcal{T})$ of the *GNS*-representation of $\mathcal{A}$ with respect to the trace $\mathcal{T}$ can be written as the direct **P**-integral of $L^2(G)$. We will not give details. A consequence of this result is that the family $\{\pi_\omega;\omega\in\Omega\}$ of representations of $\mathcal{A}$ extends to a (faithful) family of (weakly measurable) representations of $\mathcal{W}$. In particular, any spectral projection P of a Hamiltonian $H\in\mathcal{A}$ can be seen as a covariant **P**-essentially bounded family P_ω of projections, where P_ω is an eigenprojection of H_ω.

The differential structure is obtained through the derivations defined on $\mathcal{A}_0$ by

$$\partial_j A(\omega,\mathbf{x})=\imath x_j A(\omega,\mathbf{x}). \tag{33}$$

It is a *-derivation in the sense given in Sec. III A. By exponentiation it defines a two-parameter group of *-automorphisms given by

$$\rho_{\mathbf{k}}(A)(\omega,\mathbf{x})=e^{\imath\mathbf{k}\cdot\mathbf{x}}A(\omega,\mathbf{x}),$$

where $\mathbf{k}\in\mathbf{R}^2$ and $\mathbf{k}\cdot\mathbf{x}=k_1x_1+k_2x_2$. We notice that in the discrete case, $\mathbf{k}$ is defined modulo $2\pi\mathbf{Z}^2$. Introducing on $L^2(G)$ the position operator $\mathbf{X}=(X_1,X_2)$, namely, the multiplication by $\mathbf{x}$, we get

$$\pi_\omega(\rho_{\mathbf{k}}(A))=e^{-\imath\mathbf{k}\cdot\mathbf{X}}\pi_\omega(A)e^{\imath k\cdot\mathbf{X}},\quad \pi_\omega(\partial_j A)=-\imath[X_j,\pi_\omega(A)]. \tag{34}$$

We remark that in the case of periodic media, this derivation is just differentiation in quasi-momentum space. We will denote by $\mathcal{C}^N(\mathcal{A})$ the set of elements $A\in\mathcal{A}$ for which the map $\mathbf{k}\in\mathbf{R}^2\mapsto\rho_{\mathbf{k}}(A)\in\mathcal{A}$ is N-times continuously differentiable. If N is an integer, this is equivalent to say that $\|\partial_1^a\partial_2^b A\|<\infty$ for any pair a,b of integers such that $a+b\leqslant N$. In much the same way noncommutative Sobolev spaces can be defined. For the purpose of this work we will use the Sobolev space $\mathcal{S}=\mathcal{H}^2(\mathcal{A},\mathcal{T})$, namely, the Hilbert space obtained by completion of $\mathcal{A}_0$ under the Hilbert norm given by the inner product

$$\langle A|B\rangle_{\mathcal{H}^2}=\mathcal{T}(A^*B)+\mathcal{T}(\nabla A^*\cdot\nabla B),\quad A,B\in\mathcal{A}_0,$$

where $\nabla=(\partial_1,\partial_2)$ is the noncommutative gradient operator.

We will finish this section by giving a technical result which will be used later on.

Proposition 2: Let H be a self-adjoint element in $\mathcal{C}^1(\mathcal{A})$ where $\mathcal{A}=C^*(\Omega\times\mathbf{Z}^2;\mathcal{B})$ and let $\mathcal{N}$ be its density of states. Then for any function $f\in L^1(\mathbf{R},d\mathcal{N})$ we have

$$\mathcal{T}(f(H)\partial_j H)=0.$$

Sketch of the proof. By density, it is enough to prove this result for a smooth function f on $\mathbf{R}$ with compact support. Then one can write

$$f(H)=\int_{\mathbf{R}} dt\; e^{-\imath tH}\tilde{f}(t),$$

where $\tilde{f}$ is the Fourier transform of f. Since H is bounded, this integral converges in norm. Classical results on Fourier transform show that $\tilde{f}$ is a smooth rapidly decreasing function over $\mathbf{R}$. Thus, it is sufficient to show the result for $f(E)=e^{-\imath tE}$. In this case, Duhamel's formula gives[67]

$$\nabla e^{-\imath tH} = -\imath \int_0^t ds\ e^{-\imath(t-s)H} \nabla H e^{-\imath sH}.$$

Taking the trace, the left hand side vanishes, since the trace is $\rho_{\mathbf{k}}$-invariant. The right hand side is given by $\mathcal{T}(e^{-\imath tH}\nabla H)$. The proof is then complete. □

In the continuum case, a similar result holds, but it is technically more complicated. We will restrict ourselves to the case of Schrödinger operators.

Proposition 3: Let H be given by Eq. (26) with $V \in L^\infty(\mathbf{R}^2)$. We denote by Ω the hull of V. Then ∇H is a well-defined self-adjoint operator bounded from above by $(H+b)^{1/2}$ for some positive b. Moreover, if $\mathcal{N}$ denotes the DOS of H, for every $\mathcal{N}$-measurable function f such that the map $E \mapsto Ef(E)$ is in $L^1(\mathbf{R}, d\mathcal{N})$, we have

$$\mathcal{T}(f(H)\nabla H) = 0.$$

Sketch of the proof. We use the results of Ref. 61, Sec. 2.5 to conclude that e^{-tH} admits an integral kernel $F(t;\omega,\mathbf{x})$ for $\Re e(t) > 0$ because H is bounded from below. This kernel is jointly continuous in $\omega,\mathbf{x}$ and holomorphic in t in this domain. Moreover, it decays rapidly at infinity in $\mathbf{x}$ (uniformly on compact subsets for the other variables). Then we use the same argument as in Proposition 2 provided we replace $e^{-\imath tH}$ by $e^{-(\epsilon+\imath t)H}$ for any $\epsilon > 0$. □

IV. THE KUBO–CHERN FORMULA

This chapter is devoted to the Kubo formula and its relation to the Chern character. The first three sections are devoted to transport theory and are not treated on a completely rigorous footing. This is because technical proofs would require too many details and not shed more light on the physics.

A. The relaxation time approximation

The theory of transport is an essential tool in dealing with electronic properties of solids. There are numerous theoretical approaches with complexity varying from the Drude–Sommerfeld theory to the N-body framework. Whatever the starting point, the Greenwood–Kubo formulas for the transport coefficients, such as the electric or thermal conductivity, are the main consequences. They are widely used and accepted because of their correspondence with experimental results. Still their derivation from first principles is questionable. One does not really know the precise domain of validity of the linear response approximation. The occurrence of many new devices in solid state physics liable to test these ideas, such as mesoscopic systems, has raised this question again.

It is not the purpose of this work to investigate that problem. However, we have seen in Sec. II A that in the classical Hall effect, the relation between the Hall current and the Hall voltage is linear although there is no dissipation mechanism. For this reason, one might expect that the derivation of Kubo's formula for the Hall conductance in a quantum system should not require a many particle theory. This is actually not true. We will see in this section that such a point of view is very naïve and gives rise to paradoxa. Without dissipation, a quantum theory leads either to a vanishing conductivity or to an infinite one in most physically sound situations.

Moreover the great accuracy of the IQHE has been used in metrology for the definition of a new standard of resistance.[3] It is thus necessary to derive a formula allowing the control of deviations from the ideal QHE. We have already indicated in Sec. III B what the physical conditions are under which the IQHE is exact, namely under which it can be stated as a theorem within a well-defined mathematical framework.

This is why we propose a one-particle model including collision effects, such as interaction with phonons or other electrons, realizing thereby the so-called *relaxation time approximation*

(RTA). We derive a Kubo formula for the conductivity tensor which allows the linear response approach to be justified and gives control on the order of magnitude of the deviation from the ideal theory. The construction of this model is based upon the phenomenological approach which can be found in standard books in solid state physics such as Ref. 18, Chaps. 1, 2, 12, 13.

We consider the electron fluid of our system as a gas of independent fermions. Neglecting interactions between electrons is actually very rough and in many cases completely wrong. Nevertheless the Landau theory of Fermi liquids shows that such an approximation is quite acceptable if the particles are actually quasiparticles "dressed" by the interactions.[68] In particular, their mass need not necessarily be the electron mass. Moreover their lifetime is finite, but it goes to infinity as their energy gets closer to the Fermi level. Quasiparticles carrying current are therefore stable at zero temperature. In the sequel, we continue to talk of electrons having in mind that we are actually treating quasiparticles.

The advantage of this independent-electron approximation is that one can avoid using second quantization and can restrict oneself to a one-particle description. The constraint given by Pauli's principle is then represented by the use of the Fermi–Dirac distribution function.

For this reason, our starting point will be the one-particle Hamiltonian H of the form already described in Sec. III E. It includes kinetic energy of the electrons as well as whatever time-independent forces there are acting on them in the crystal. These latter include the periodic potential created by the ions and the aperiodic corrections due to impurities, defects, and other kinds of deformations. This Hamiltonian goes beyond the band theory since it may include Anderson localization for instance. We will assume that H is affiliated to the observable algebra $\mathcal{A}=C^*(\Omega,G,\mathcal{B})$ where $G=\mathbf{R}^2$ or $\mathbf{Z}^2$ and Ω is some compact space (see Sec. III F). In the grand canonical ensemble the thermal equilibrium at inverse temperature $\beta=1/k_BT$ (where k_B is the Boltzmann constant and T is the temperature) and chemical potential μ is described by means of the Fermi–Dirac distribution. In the algebraic language it means that, if $A\in\mathcal{A}$ is a one-particle extensive observable, its thermal average per volume is given by

$$\langle A\rangle_{\beta,\mu}=\mathcal{T}(Af_{\beta,\mu}(H)),\quad \text{with}\quad f_{\beta,\mu}(H)=(1+e^{\beta(H-\mu)})^{-1}. \tag{35}$$

One important example for an observable is the current operator

$$\mathbf{J}_\omega=q\,\frac{d\mathbf{X}}{dt}=\frac{\imath q}{\hbar}\,[H_\omega,\,\mathbf{X}].$$

Using the differential structure (33), this can be written as

$$\mathbf{J}_\omega=\frac{q}{\hbar}(\nabla H)_\omega.$$

It is physically obvious that the average current vanishes since the velocity distribution is usually symmetric under the change of its sign. This is actually what happens within our framework because of the Propositions 2 and 3 and the formula (35) above. Producing a nonvanishing average current requires imposing an external force such as an electromagnetic field. Let us consider the simplest case in which we superimpose a constant uniform electric field $\mathcal{E}$ on our system for time $t\geqslant 0$. Then the evolution of an observable is no longer given by H_ω but rather by $H_{\omega,\mathcal{E}}=H_\omega-q\mathcal{E}\cdot\mathbf{X}$. We notice that while H is affiliated to $\mathcal{A}$, this is certainly not true for this Hamiltonian since the position operator is not homogeneous. However, the evolution under this operator leaves $\mathcal{A}$ invariant. For indeed, whenever A is smooth in $\mathcal{A}$, Heisenberg's equation reads

$$\frac{dA_\omega}{dt}=\frac{\imath}{\hbar}\,[H_\omega,A_\omega]+\frac{q\mathcal{E}}{\hbar}\cdot(\nabla A)_\omega. \tag{36}$$

If in particular H belongs to $\mathscr{A}$, the right-hand side of (36) stays in $\mathscr{A}$ so that we expect this equation to be integrable within $\mathscr{A}$. This can be proved by use of the Trotter product formula.[69] We denote by η_t the evolution given by H, namely,

$$\eta_t(A)=e^{i(tH/\hbar)}Ae^{-i(tH/\hbar)}. \tag{37}$$

Since H is affiliated to $\mathscr{A}$, this evolution leaves $\mathscr{A}$ invariant and defines a one-parameter group of *-automorphisms of $\mathscr{A}$. Because of formula (34), we also have

$$\pi_\omega(\rho_{(qt/\hbar)\mathscr{E}}(A))=e^{-i(qt\mathscr{E}\cdot\mathbf{X}/\hbar)}\pi_\omega(A)e^{i(qt\mathscr{E}\cdot\mathbf{X}/\hbar)}.$$

This last evolution also leaves $\mathscr{A}$ invariant. By means of the Trotter product formula we find

$$e^{i(t/\hbar)\,H_{\omega,\mathscr{E}}}\pi_\omega(A)e^{-i(t/\hbar)\,H_{\omega,\mathscr{E}}}=s-\lim_{N\to\infty}\pi_\omega((\eta_{\frac{t}{N}}\rho_{qt\mathscr{E}/N\hbar})^N(A)).$$

Here, s-lim is the limit in the strong operator topology. So this defines a one-parameter group of automorphisms $\eta_t^{\mathscr{E}}$ of the von Neumann algebra $\mathscr{W}$. It represents the evolution of the observables after the electric field has been turned on. We will not investigate here whether this group of automorphisms leaves $\mathscr{A}$ itself invariant. Let us only notice that, whenever H is bounded, a Dyson expansion shows that $\eta_t^{\mathscr{E}}$ is an automorphism of $\mathscr{A}$. Moreover, we only have to work in the Hilbert space $L^2(\mathscr{A},\mathscr{T})$ so that the previous result will be sufficient if H is unbounded. The new current at time $t\geqslant 0$ is then formally given by

$$\mathbf{J}(t)=\eta_t^{\mathscr{E}}(\mathbf{J}).$$

Since the Hamiltonian $H_{\omega,\mathscr{E}}$ no longer commutes with H_ω, the thermal average $\mathbf{j}(t)=\langle\mathbf{J}(t)\rangle_{\beta,\mu}$ of the current will no longer vanish in general. The macroscopic response we expect from a constant uniform electric field is a constant and uniform current, but the microscopic quantal forces lead to fluctuations of $\mathbf{j}(t)$ in time. To extract the response, it is therefore necessary to consider the time average of this current which is actually what one measures in experiment. For indeed, the typical relaxation time of an apparatus measuring the current is of the order of 1 ms to 1 μs (unless very short times are needed), and this has to be compared with the typical collision time of the order of 10^{-13} s for the best conductors. Thus, we set

$$\mathbf{j}_{\beta,\mu,\mathscr{E}}=\lim_{T\to\infty}\frac{1}{T}\int_0^T dt\;\mathbf{j}(t)=\lim_{\delta\downarrow 0}\;\delta\int_0^\infty dt\;e^{-t\delta}\mathbf{j}(t). \tag{38}$$

However, our model does not take collisions into account. Indeed we have

Proposition 4: If the Hamiltonian H is bounded in $\mathscr{A}$, the projection of the time and thermal averaged current $\mathbf{j}_{\beta,\mu,\mathscr{E}}$ along the direction of the electric field $\mathscr{E}$ vanishes.

Proof. Let us compute $\mathscr{E}\cdot\mathbf{J}(t)$. Using the Heisenberg equation it is easy to see that

$$\mathscr{E}\cdot\mathbf{J}(t)=\frac{dH(t)}{dt},\quad H(t)=\eta_t^{\mathscr{E}}(H).$$

Taking the time average gives us

$$\frac{1}{T}\int_0^T dt\,\mathscr{E}\cdot\mathbf{J}(t)=\frac{H(T)-H}{T}.$$

Since H is bounded in norm and since $\|H(t)\|=\|H\|$, the right hand side vanishes as $T\to\infty$. □

This result is easy to understand in the one-band approximation. In this case, the electric field produces a time shift of the quasimomentum, namely, $\mathbf{k}\in\mathbf{B}\mapsto\mathbf{k}(t)=\mathbf{k}-q\mathscr{E}t/\hbar$. Taking the time average is therefore equivalent to averaging over quasimomenta, and this is exactly zero. We point out that this result applies also to models with disorder of the form given in (28). The Anderson model is a special example of such Hamiltonians. More generally, if we accept that transitions to bands far from the Fermi level are essentially forbidden, the effective Hamiltonian is always bounded and the Proposition 4 then leads to the vanishing of the current component parallel to $\mathscr{E}$.

For these reasons, the presented approach is definitely too naïve. As we already pointed out, collisions occur after time periods very short compared to the time over which we average the current. These collisions produce a loss of memory in the current evolution and are actually responsible for the nonvanishing of the time average. Everything looks like as if time evolution has to be considered over short intervals only. We propose to take collision effects into account by means of the following model: the time evolution is described by the new time-dependent Hamiltonian $H_{coll}(t)=H-q\mathscr{E}\cdot\mathbf{X}+W_{coll}(t)$ where

$$W_{coll}(t)=\sum_{n\in\mathbf{Z}}W_n\delta(t-t_n).$$

In this expression, the t_n's represent collision times. They are labeled such that $\ldots,t_{-1}<t_0=0<t_1<\cdots<t_n<t_{n+1}<\cdots$. Because these times occur randomly, we will assume that the time delays $\tau_n=t_n-t_{n-1}$ are independent identically distributed random variables distributed according to Poisson's law with mean value $\mathbf{E}(\tau_n)=\tau$.

The W_n's are the collision operators. Their main effect is to produce a loss of memory during the time evolution of the current. This process should enforce thermal equilibrium. In particular, it should not modify the Fermi–Dirac distribution. The only way to respect this constraint is to force the W_n's to commute with the Hamiltonian H, but to be random otherwise. More precisely, we will assume that the W_n's are independent identically distributed random operators, commuting with the Hamiltonian H. Their distribution is supposed to be symmetric under the change of sign $W_n\leftrightarrow -W_n$. Let then $\hat{\kappa}$ be the operator acting on $\mathscr{A}$ as

$$\hat{\kappa}(A)=\mathbf{E}(e^{iW_n/\hbar}Ae^{-iW_n/\hbar}).$$

It can be extended to a bounded operator on $L^2(\mathscr{A},\mathscr{T})$. Moreover, because of the change-of sign symmetry, it is self-adjoint. Let then $\mathscr{A}_H$ be the closed subspace of $L^2(\mathscr{A},\mathscr{T})$ generated by bounded functions of H. We will then assume that there is $0\leq\kappa<1$ such that

$$\|\hat{\kappa}(A)\|\leq\kappa\|A\|,\quad\forall A\in(\mathscr{A}_H)^{\perp}.$$

Since W_n commutes with H it follows that $\hat{\kappa}$ leaves $\mathscr{A}_H$ invariant so that it forces any operator $A\in\mathscr{A}$ along the direction of H. We will not give explicit examples of such random operators because we will only use the *collision efficiency operator* $\hat{\kappa}$ later on, so that it solely characterizes our model. The parameter κ is purely phenomenological and represents an *efficiency coefficient* of the collision process. The smaller κ, the more efficient are the collisions. We see that the relaxation time, namely, the time after which there is a complete loss of memory of the initial evolution, has to be renormalized by the efficiency coefficient. We will see below that $\tau_{rel}=\tau/(1-\kappa)$ is a good measure for this relaxation time. We will discuss later on how to choose this parameter to fit with real systems.

B. Kubo's formula

We now follow the strategy defined in the previous section and compute the current average with the collisions taken into account. This requires to calculate the evolution operator $S_\xi(t)$ between time 0 and $t \geq 0$, where ξ represents the random variables $\xi=(\tau_n, W_n)_{n>0}$. It is well known that a kick term like $W_n\delta(t-t_n)$ produces a contribution $e^{\imath W_n/\hbar}$ in the evolution between times t_n-0 and t_n+0, namely, at the kick time.[70,71] Therefore, if $n \geq 1$ and $t_{n-1} \leq t < t_n$, we find

$$S_\xi(t)=e^{\imath(t-t_{n-1})H/\hbar}\prod_{j=1}^{n-1} e^{\imath W_j/\hbar}e^{\imath(t_j-t_{j-1})H/\hbar}.$$

We will now set $\mathscr{L}_H(A)=(\imath/\hbar)[H, A]$ for A in a suitable dense subalgebra of $\mathscr{A}$. This is a *-derivation of $\mathscr{A}$ because of (37). Therefore, it defines an anti-self-adjoint operator on $L^2(\mathscr{A},\mathscr{T})$. Moreover, the evolution of observables is given by the automorphism

$$\eta^{\mathscr{E}}_{\xi,t}=e^{(t-t_{n-1})(\mathscr{L}_H+(q/\hbar)\mathscr{E}.\nabla)}\prod_{j=1}^{n-1} e^{\mathscr{L}_{W_j}}e^{(t_j-t_{j-1})(\mathscr{L}_H+(q/\hbar)\mathscr{E}.\nabla)},$$

where $\mathscr{L}_{W_j}(A)=(\imath/\hbar)[W_j, A]$. The operator $\eta^{\mathscr{E}}_{\xi,t}$ may also be seen as a unitary on $L^2(\mathscr{A},\mathscr{T})$. In view of Eq. (38), the time average of this evolution requires the calculation of

$$\hat{\eta}^{\mathscr{E}}_\delta=\delta\int_0^\infty dt\, e^{-t\delta}\mathbf{E}(\eta^{\mathscr{E}}_{\xi,t}),$$

where $\mathbf{E}$ denotes the average over ξ and $\delta>0$. After a straightforward calculation we find

$$\hat{\eta}^{\mathscr{E}}_\delta=\delta\left(\delta+\frac{1-\hat{\kappa}}{\tau}-\left(\mathscr{L}_H+\frac{q}{\hbar}\mathscr{E}.\nabla\right)\right)^{-1}.$$

This expression is meaningful because the operator in parentheses has a real part bounded from below by δ. The average current is then given by

$$\mathbf{j}_{\beta,\mu,\mathscr{E}}(\delta)=\frac{q}{\hbar}\mathscr{T}(f_{\beta,\mu}(H)\,\hat{\eta}^{\mathscr{E}}_\delta(\nabla H)). \tag{39}$$

We then remark that, for $\mathscr{E}=0$, the right-hand side of (39) vanishes because $\hat{\kappa}(f(H))=f(H)$ and $\mathscr{L}_H(f(H))=0$ for any function f, whereas $\mathscr{L}_H$ is anti-self-adjoint. Subtracting this expression with $\mathscr{E}=0$ and using the inner product $\langle A|B\rangle=\mathscr{T}(A^*B)$ in $L^2(\mathscr{A},\mathscr{T})$, gives

$$\mathbf{j}_{\beta,\mu,\mathscr{E}}(\delta)=\left(\frac{q}{\hbar}\right)^2\sum_{i=1,2}\mathscr{E}_i\left\langle \partial_i f_{\beta,\mu}(H)\left|\frac{1}{\delta+(1-\hat{\kappa})/\tau-\mathscr{L}_H-(q/\hbar)\mathscr{E}.\nabla}\nabla H\right.\right\rangle. \tag{40}$$

If we assume that the nonlinear term in the electric field is negligible, we can let δ converge to zero. For indeed, ∇H and $\nabla f_{\beta,\mu}(H)$ are orthogonal to $\mathscr{A}_H$. Thus it is enough to consider the restriction of the operator $\delta+(1-\hat{\kappa})/\tau-\mathscr{L}_H$ to the subspace orthogonal to $\mathscr{A}_H$. This restriction has a real part bounded from below by $(1-\kappa)/\tau>0$. Thus it is an invertible operator. If we do not neglect the nonlinear term, we have to investigate more thoroughly what happens as $\delta\downarrow 0$. We will not discuss that matter here and postpone it to a future work. Letting δ converge to zero, we will get the Kubo formula for the conductivity tensor:

$$\mathbf{j}_{\beta,\mu,\kappa}=\boldsymbol{\sigma}\mathcal{E}, \quad \text{with} \quad \sigma_{ij}=\left(\frac{q^2}{\hbar}\right)\left\langle \partial_j f_{\beta,\mu}(H)\left|\frac{1}{\hbar\{(1-\hat{\kappa})/\tau-\hbar\mathscr{L}_H}\,\partial_i H\right\rangle\right. . \tag{41}$$

Let us remark that the quantity in the bracket is dimensionless in $2D$: the trace $\mathscr{T}$ is a trace per unit volume (so that it has the dimension of the inverse of an area) while the derivative ∂_j has the dimension of a length. In addition $\hbar\mathscr{L}_H$ has the dimension of an energy. Therefore, we get σ dimensionless in units of q^2/h.

If, in addition, the system is symmetric under rotation by $\pi/2$ in space, namely, if the Hamiltonian H and the efficiency operator $\hat{\kappa}$ are invariant under such a rotation, then the conductivity tensor can be written in matrix form as

$$\sigma=\begin{pmatrix} \sigma_{//} & \sigma_H \\ -\sigma_H & \sigma_{//} \end{pmatrix},$$

where the Hall conductance is the off-diagonal term σ_H, while $\sigma_{//}$ is the direct conductivity.

Let us now consider the limit for which the IQHE is valid. Namely, the electric field is vanishingly small, the temperature is zero, and the relaxation time is infinite. In this limit, the Fermi distribution becomes

$$\lim_{\beta\uparrow\infty} f_{\beta,\mu}(H)=P_F, \tag{42}$$

where P_F, the *Fermi projection*, is the spectral projection onto energy levels lower than the Fermi energy. Here the limit is taken with respect to the norm in $L^2(\mathscr{A},\mathscr{T})$. Actually, this result is correct only if the Fermi level E_F is a continuity point of the DOS of H. Otherwise the eigenprojection $P_{\{E_F\}}$ corresponding to the eigenvalue E_F satisfies $\mathscr{T}(P_{\{E_F\}})>0$ and therefore defines a nonzero element of $L^2(\mathscr{A},\mathscr{T})$. This is what happens for instance in the case of the Landau Hamiltonian whenever $E_F=(n+1/2)\hbar\omega_c$. Moreover, we need the Fermi projection to be *Sobolev differentiable*, namely it has to satisfy $\nabla P_F\in L^2(\mathscr{A},\mathscr{T})$, otherwise the formal limit is meaningless. We will see in Sec. V B below that such a condition is related to the finiteness of localization length at the Fermi level. We also need to show that the limit of the derivative of the Fermi distribution is the derivative of its limit. Even though we know this to be true for physically reasonable conditions on H, we will not give the proof here but postpone it to a future work.

On the other hand, in the limit of infinite relaxation time, we are left with the formal expression $\mathscr{L}_H^{-1}\partial_i H$ which is meaningless in general. If we formally diagonalize the Hamiltonian H, matrix elements of that expression are given by

$$\langle E|(\hbar\mathscr{L}_H)^{-1}\partial_i H|E'\rangle=\frac{\langle E|\partial_i H|E'\rangle}{E-E'}\,.$$

In particular it diverges whenever $E\approx E'$ unless the numerator vanishes for some reason. This divergency does, however, not occur in the expression of the conductivity tensor. The reason is that the derivative ∇P_F has nonvanishing matrix elements only between energies E and E' such that $E<E_F<E'$ or $E'<E_F<E$. For indeed, if P is a projection and d is a derivation, we see that

$$dP=d(P^2), \quad \Rightarrow \quad dP=(1-P)dPP+PdP(1-P).$$

So we need to consider the operators $P_F\mathscr{L}_H^{-1}\partial_i H\,(1-P_F)$ and $(1-P_F)\mathscr{L}_H^{-1}\partial_i H\,P_F$ only. We then obtain

Lemma 1: If the Fermi level is not a discontinuity point of the DOS of H, and if the Fermi projection is Sobolev differentiable, the following formulas hold

$$P_F \mathscr{L}_H^{-1} \partial_i H \ (1-P_F) = -\imath\hbar P_F \partial_i P_F (1-P_F),$$

$$(1-P_F)\mathscr{L}_H^{-1} \partial_i H \ P_F = \imath\hbar (1-P_F)\partial_i P_F P_F .$$

Proof. Let us consider the first formula only, because the other can be treated in the same way (notice however the change of sign). Let B_+ be the right-hand side. Then

$$\hbar \mathscr{L}_H(B_+) = P_F[H, \partial_i P_F](1-P_F).$$

Since H commutes with P_F, we find $[H, \partial_i P_F] = -[\partial_i H, P_F]$ (after using basic properties of derivations). This gives immediately

$$\hbar \mathscr{L}_H(B_+) = P_F \partial_i H (1-P_F).$$

Since B_+ connects only energies below the Fermi level with energies above it, and since the Fermi level is not an eigenvalue of H, $\mathscr{L}_H$ is invertible on the subspace of such operators, giving the result of the lemma. □

Corollary 2 (IQHE-Kubo Formula): If the Fermi level is not a discontinuity point of the DOS of H, in the zero temperature and infinite relaxation time limit and provided the Fermi projection is Sobolev differentiable, namely, $\nabla P_F \in L^2(\mathscr{A}, \mathscr{T})$, the conductivity tensor is given by

$$\sigma_{ij} = \frac{q^2}{h} 2\imath\pi \mathscr{T}(P_F[\partial_i P_F, \partial_j P_F]) \ .$$

In particular the direct conductivity vanishes.

C. Estimating the deviations from the IQHE limit

Before returning to the IQHE, we want to give some idea of the accuracy of the IQHE–Kubo formulas given in the Corollary 2. We recall that it is obtained under the following conditions:

(i) the area of the sample is considered as infinite;
(ii) we work within the relaxation time approximation;
(iii) the electric field is vanishingly small;
(iv) the temperature is zero;
(v) the relaxation time is infinite.

In this section we will evaluate roughly the size of the correction terms whenever one of these conditions is relaxed. We know that the relative error on the IQHE measurement of the universal constant e^2/h is of order of 2×10^{-8} at best.[20] Therefore the neglected terms should be smaller than this number in order that the experiment is reliable.

We will not estimate seriously the finite-size effects even though they are in principle accessible to a mathematical estimate within the noncommutative framework. However, it is generally accepted that these effects decrease to zero as $e^{-L/r}$ where L is the sample size and r some typical length. We will choose r to be of the order of magnitude of the localization length. We shall see that it diverges precisely whenever the Hall conductance jumps from one plateau to another. Then as this localization length increases, there is a critical value beyond which the Coulomb interaction between electrons can no longer be neglected. This is the situation in which the FQHE occurs. To estimate this value, one can define r to be such that the electrostatic energy of a pair of electrons separated by a distance r is of the order of magnitude of the cyclotron energy. This gives $r \approx 1$ μm namely, an overestimated large distance compared to the magnetic length. Thus as soon as $L \geqslant 20$

μm, the finite size effects are negligible on the integer plateaux. Fluctuations from sample to sample due to finite size and disorder will then be negligible. Let us indicate that breakdown of the IQHE due to finite size effects has been observed.[23]

The relaxation time approximation should actually be enough to estimate other effects. Corrections to such an approximation should not be effective at zero temperature, since the relaxation time depends almost only upon the energy level. Only if transitions between different bands are taken into account, it is necessary to go beyond this approximation. This problem is too difficult to be investigated here.

Let us consider the effect of a nonzero electric field. Nonlinear effects such as bistability or hysteresis have indeed been observed in such devices.[72] If formula (40) is correct, the correction term due to a nonzero electric field is given by

$$\delta\sigma_{i,j}=\left(\frac{q}{\hbar}\right)^2\left\langle \partial_j f_{\beta,\mu}(H)\left|\frac{1}{((1-\hat{\kappa})/\tau-\mathscr{L}_H}\,\frac{q\mathscr{E}.\nabla}{\hbar}\,\frac{1}{((1-\hat{\kappa})/\tau-\mathscr{L}_H}\,\partial_i H\right\rangle\right.\!.$$

The relative error is measured by the size of the ratio $q\mathscr{E}.\nabla/\hbar\mathscr{L}_H$ between the electric energy and the level separation. If we choose a level separation of the order of the distance between Landau levels, namely, $\hbar\omega_c$, and let the electric energy be of the order of $e\mathscr{E}a$ (where a is the atomic distance, namely, 1 Å), we find $\delta\sigma/\sigma\approx 10^{-7}$ for an electric field of 1 V/m. Thus nonlinear effects require higher electric fields and it is very easy to choose the electric field in such a way as to make this correction negligible.

Nonzero temperature effects at infinte relaxation time can be estimated as follows. Coming back to Eq. (42), using Eq. (41) and the Lemma 1, we can write the Kubo formula at finite temperature as

$$\sigma_H=\int_{-\infty}^{\infty} dE\, f'_{\beta,\mu}(E)\frac{q^2}{h}\,2\imath\pi\mathscr{T}(P_{\leqslant E}[\partial_1 P_{\leqslant E},\,\partial_2 P_{\leqslant E}]), \tag{43}$$

provided we assume the relaxation time to be infinite. Here $P_{\leqslant E}$ is the eigenprojection of the Hamiltonian on energies smaller than or equal to E. The spectral theorem allowed us to write (with an integration by parts):

$$f(H)=\int_{-\infty}^{\infty} dP_{\leqslant E}\, f(E)=-\int_{-\infty}^{\infty} dE\, f'(E)P_{\leqslant E},$$

whenever f is a bounded differentiable function vanishing rapidly at $+\infty$. We will show that the integrand is quantized and equal to ne^2/h whenever $\hbar\omega_c(n-1/2)<E<\hbar\omega_c(n+1/2)$ with $n=0,1,2,\ldots$. This is actually true for the most common devices. If they have charged impurities the jump does not necessarily occur at the Landau level.[25] But this is of no importance for the present discussion since we only need orders of magnitude for the correction terms. Incidentally Eq. (43) shows that the direct conductivity vanishes at finite temperature as well (if the relaxation time is infinite). The error term can easily be computed and is given by

$$\frac{\delta\sigma_H}{\sigma_H}=\sum_{n'\leqslant 0,n'\neq n}\frac{n'}{n}\int_{\hbar\omega(n'-1/2)}^{\hbar\omega(n'+1/2)} dE\frac{\beta}{4\,\cosh^2(\beta(E-\mu)/2)}\,, \tag{44}$$

provided $\hbar\omega_c(n-1/2)<E_F<\hbar\omega_c(n+1/2)$. We know that the chemical potential equals E_F up to small terms of order T^2 (where T is the temperature). If E_F is located at a distance $r\hbar\omega_c$ from the nearest Landau level, where $0<r\leqslant 1/2$, this correction is thus of the order of $e^{-r\beta\hbar\omega_c}$. For a magnetic field of $10T$ and a charge carrier mass of $0.1m_e$, $\hbar\omega_c/k_B\approx 140$ °K. This term is smaller than 10^{-7} for temperature lower than 4 °K. Accurate experiments are often performed at 50 mK

and magnetic fields can be as high as $18T$. Moreover, in heterojunctions, the effective mass of the charge carriers is one order of magnitude smaller. Therefore, the pure temperature effect (at infinite relaxation time) is indeed negligible.

The infinite relaxation time approximation is in fact the most important effect. The relevant parameter is the relaxation time $\tau_{rel}=\tau/(1-\kappa)$; it can be estimated by means of Drude's formula by measuring conductivities. Assuming the relaxation time to be large, the lowest order contribution to the Hall conductivity is given by [see Eq. (41)]

$$\sigma_H \approx \frac{q^2}{\hbar}\,\frac{1}{\tau_{rel}}\,\mathcal{T}(P_F[\mathcal{L}_H^{-1}(\partial_1 P_F),\ \partial_2 P_F]). \tag{45}$$

Estimating $\mathcal{L}_H^{-1}$ by $1/\omega_c$, we see that the error term due to this contribution is controlled by $\varepsilon=(\tau\omega_c)^{-1}$. For semiconductors used in the IQHE, the mobility μ_c (c denotes the type of charge carriers) at zero magnetic field and 4.2 °K varies from about 10^4 cm^2/V s for the MOSFET to 10^6 cm^2/V s for the AlGaAs or InAs-GaSb heterojunctions (Ref. 20, pp. 40 and 41). These high-mobility devices are used for the FQHE. On the other hand, the effective mass of charge carriers varies from $0.2m_e$ for the MOSFET to $0.03m_e$ for InAs-GaSb heterojunctions. Since $\tau\omega_c=\mu_c\mathcal{B}$, we find $\varepsilon\approx 0.1\sim 0.005$ for $\mathcal{B}=10T$. We remark, however, that our relaxation time τ_{rel} only incorporates time-dependent disorder and may therefore be significantly smaller than the one calculated with Drude's formula.

It is usually estimated[2] that the deviation of the Hall conductance from its ideal value is linked to the direct conductivity $\sigma_{//}$ by

$$\delta\sigma_H \approx \frac{\sigma_{//}}{\tau_{rel}\omega_c}\,.$$

Measurements give $\min(\sigma_{//}) \leqslant 10^{-7}$ $\max(\sigma_{//}) \approx 10^{-7} e^2/h$ (Ref. 2) so that the relative error on the Hall conductance is indeed of the order of 10^{-8}.

Why is $\sigma_{//}$ so small? Looking at the equivalent of Eq. (45) for the direct conductivity, we see that one way to estimate it consists of replacing $\mathcal{L}_H^{-1}$ by $1/\omega_c$ whereas the remaining terms are related to the localization length λ. But since $\mathcal{T}(P_F)=n$ is the charge carrier density, we expect this term to be of the order of $n\lambda^2/\tau_{rel}\omega_c$. The mobility μ_c of the charge carriers c is $\tau_{rel}\omega_c/\mathcal{B}$. Moreover we introduce the filling factor $\nu=nh/q\mathcal{B}$ leading to the very rough estimate

$$\frac{\delta\sigma_H}{\sigma_H} \leqslant \text{const.}\,\nu\,\frac{e}{h}\,\frac{\lambda^2}{\mu_c}\,.$$

For the heterojunction, ν is typically between 1 and 10 at most. The ratio q/h is given by the electron charge e/h and is thus universal. The localization length is always bigger than 80 Å. The mobility is at most equal to $2\cdot 10^6$ cm^2/V s for the best heterojunctions. We then find a ratio of the order of 10^{-4}. Therefore, this estimate is too crude to explain the high accuracy observed in experiment. Nevertheless, it shows that collision effects are dominant whereas localization is a necessity.

In any way, physical arguments indicate[73,20] that because of the small value of the density of states between Landau levels, the leading contribution to the direct conductivity is given by phonon-assisted hopping, at least while the Fermi energy is not too close to a Landau level. Estimation (44) only includes conductivity by thermal activation, whereas (45) only incorporates effects due to a finite collision rate. Here, we shall not treat the interplay of the two phenomena, but only indicate that Mott theory leads to $\sigma_{//}\approx\sigma_0\exp(-(T_0/T)^{1/3})$ in two dimensions, while including Coulomb interaction would give $\sigma_{//}\approx\sigma_0\exp(-(T_0/T)^{1/2})$;[74] this latter is in better agreement with experiment.[75]

D. Dixmier trace and Sobolev space

In this section, we introduce the Dixmier trace and we will prove a formula that can be found in a similar form in a paper by Connes.[60] We will see that this tool is a key point both in proving the integrality of the Chern number of a projection and in the study of localization. Most of the material presented here can be found in Refs. 6 and 60 so that we will give no proof unless absolutely necessary.

First, let us recall that, given a separable Hilbert space $\mathscr{H}$, $\mathscr{K}(\mathscr{H})$ is the C^*-algebra of compact operators on $\mathscr{H}$, namely, the norm-closure of the set of finite rank linear operators on $\mathscr{H}$. The Schatten ideal $\mathscr{L}^p(\mathscr{H})$ is defined as the set of compact operators T such that $\Sigma_{n=1}^{\infty}\mu_n^p<\infty$, where the μ_n's are the eigenvalues of $(TT^*)^{1/2}$ labeled in the decreasing order. The following proposition summarizes the main properties of Schatten ideals; it can be found in,[76,5] for example.

Proposition 5: Let $\mathscr{L}(\mathscr{H})$ be the algebra of bounded operators on $\mathscr{H}$ and Tr *the usual trace on $\mathscr{L}(\mathscr{H})$. Then we have the following:*

(i) $\mathscr{L}^p(\mathscr{H})=\{T\in\mathscr{L}(\mathscr{H})\,|\,\mathrm{Tr}(|T|^p)<\infty\}$.
(ii) *$\mathscr{L}^p(\mathscr{H})$ is a two-sided ideal in $\mathscr{L}(\mathscr{H})$.*
(iii) *$\mathscr{L}^p(\mathscr{H})$ is a Banach space for the norm* $\|T\|_p=(\mathrm{Tr}(|T|^p))^{1/p}=(\Sigma_n(\mu_n(T)^p))^{1/p}$.
(iv) $\mathscr{L}^p(\mathscr{H})\subset\mathscr{L}^q(\mathscr{H})$ *for* $p\leq q$.
(v) *Let* $p,q,s\in[1,\infty)$ *with* $1/r=1/p+1/q$, $S\in\mathscr{L}^p(\mathscr{H})$ *and* $T\in\mathscr{L}^p(\mathscr{H})$. *Then, Hölder's inequality holds:* $\|ST\|_r\leq\|S\|_p\|T\|_q$.

Now we introduce the Mačaev ideals $\mathscr{L}^{p+}(\mathscr{H})$ and $\mathscr{L}^{p-}(\mathscr{H})$ and the Dixmier trace.

Definition 4: Let $\mathscr{H}$ be a separable Hilbert space and $\mathscr{K}$ be the ideal of compact operators on $\mathscr{H}$. For $p\in[1,\infty)$, the Mačaev ideal $\mathscr{L}^{p+}(\mathscr{H})\subset\mathscr{K}$ is the set of compact operators T of which the characteristic values satisfy

$$\limsup_{N\to\infty}\frac{1}{\ln N}\sum_{n=1}^{N}\mu_n^p<\infty,$$

where the characteristic values are the eigenvalues (μ_n) of $(TT^*)^{1/2}$ labeled in decreasing order. $\mathscr{L}^{p-}(\mathscr{H})$ is defined in the much same way but with the lim sup *equal to zero. We will also set*

$$\|T\|_{p+}=\sup_{N\to\infty}\frac{1}{\ln N}\sum_{n=1}^{N}\mu_n^p.$$

Theorem 7:

(i) $\mathscr{L}^{p+}$ and $\mathscr{L}^{p-}$ are two-sided ideals in $\mathscr{L}(\mathscr{H})$.
(ii) For $p\in[1,\infty)$ one has $\mathscr{L}^p\subset\mathscr{L}^{p-}\subset\mathscr{L}^{p+}\subset\mathscr{L}^{p+\varepsilon}\forall\varepsilon>0$.
(iii) The expression $\|T\|_{p+}$ defines a norm on $\mathscr{L}^{p+}$, making it into a Banach space.

Next, the Dixmier trace is constructed as follows.[59] Let Lim be a positive linear functional on the space of bounded sequences $l_+^{\infty}(\mathbf{N})$ of positive real numbers which is translation and scale invariant. If $\alpha\in l_+^{\infty}(\mathbf{N})$ converges, then the functional Lim satisfies:

$$\mathrm{Lim}(\alpha)=\lim_{n\to\infty}\alpha_n. \tag{46}$$

Scale invariance means that $\mathrm{Lim}(\alpha)=\mathrm{Lim}(\alpha_1,\alpha_1,\alpha_2,\alpha_2,\ldots)$. To construct Lim, Dixmier uses an invariant mean on the Euclidean group of $\mathbf{R}$ (the existence of such means results from a theorem of von Neumann).

Definition 5: For positive $T\in\mathscr{L}^{1+}$ and a fixed scale-invariant and positive linear functional Lim *on $l_+^\infty(\mathbf{N})$ satisfying (46), the Dixmier trace is defined by*

$$\mathrm{Tr}_{\mathrm{Dix}}(T)=\mathrm{Lim}\left(\frac{1}{\ln N}\sum_{n=1}^{N}\mu_n\right).$$

Remark that $T\in\mathscr{L}^{1+}$ if and only if $\mathrm{Tr}_{\mathrm{Dix}}(|T|)<\infty$. Moreover, if the sequence $((1/\ln N)\,\Sigma_{n=1}^{N}\mu_n)$ converges, then all functionals Lim of the sequence are equal to the limit and the Dixmier trace is given by this limit. From this definition, one can show that $\mathrm{Tr}_{\mathrm{Dix}}$ is a trace in the following sense:[59]

Proposition 6: The functional $\mathrm{Tr}_{\mathrm{Dix}}$ *defined in Definition 5 can be extended as a linear form on $\mathscr{L}^{1+}$ such that:*

(i) positivity: *if $T\in\mathscr{L}^{1+}$ is a positive operator, then* $\mathrm{Tr}_{\mathrm{Dix}}(T)>0$,
(ii) trace property: *if $S,T\in\mathscr{L}^{1+}$ then* $\mathrm{Tr}_{\mathrm{Dix}}(ST)=\mathrm{Tr}_{\mathrm{Dix}}(TS)$,
(iii) unitary invariance: *if $T\in\mathscr{L}^{1+}$ and U is unitary then* $\mathrm{Tr}_{\mathrm{Dix}}(UTU^{-1})=\mathrm{Tr}_{\mathrm{Dix}}(T)$,
(iv) continuity: *it is continuous with respect to the seminorm $\|T\|_{1+}$. Moreover,* $\mathrm{Tr}_{\mathrm{Dix}}$ *vanishes on L^{1-}.*

Let us introduce the operator $\hat{\delta}$ acting on a linear operator A on $l^2(G)$ as

$$\hat{\delta}A=[u,A],\quad u=\frac{X_1+\imath X_2}{|X_1+\imath X_2|},$$

where X_1,X_2 are the components of the position operator. The main result of this section is given by the following proposition.

Proposition 7: Let Ω be a compact metrizable space on which G acts by homeomorphisms. Let $\mathbf{P}$ *be a G-invariant ergodic probability on Ω. One then denotes by $\mathscr{A}$ the C*-algebra of this dynamical system and by $\mathscr{T}$ the trace on $\mathscr{A}$ corresponding to* $\mathbf{P}$. *Let $\mathscr{A}_0$ be the dense subalgebra of continuous functions with compact support on $\Omega\times G$.*

Then for every $A\in\mathscr{A}_0$, the following formula holds:

$$\mathscr{T}(|\nabla A|^2)=\frac{2}{\pi}\,\mathrm{Tr}_{\mathrm{Dix}}(|\hat{\delta}A_\omega|^2),\quad \text{for } \mathbf{P}\text{-almost all } \omega. \tag{47}$$

If $\mathscr{S}$ denotes the Sobolev space associated to $\mathscr{T}$, this formula can be continued to elements $A\in\mathscr{S}$. In particular, if $A\in\mathscr{S}$, then $\hat{\delta}A_\omega\in\mathscr{L}^{2+}$ for $\mathbf{P}$*-almost all ω.*

In the remaining part of this section we present the proof of Proposition 7 for the case of discrete physical space $G=\mathbf{Z}^2$; the continuous case will be treated in future work.[64] The first step in this proof is the following lemma:

Lemma 2: Let T be a bounded operator on $\ell^2(\mathbf{Z}^D)$ such that

(i) *$\exists r$ such that $\langle n|T|m\rangle=0\ \forall|n-m|\geqslant r$.*
(ii) *There exists a positive constant C such that $|\langle n|T|m\rangle|\leqslant C/(1+|n|^D)\forall m\in\mathbf{Z}^D$.*

T is then in the Mačaev ideal $\mathscr{L}^{1+}$ and, for any linear functional Lim, *its Dixmier trace can be calculated by*

$$\mathrm{Tr}_{\mathrm{Dix}}(T)=\mathrm{Tr}_{\mathrm{Dix}}(\mathrm{Diag}(T)),$$

where $\mathrm{Diag}(T)$ *is the diagonal matrix such that* $\langle n|\mathrm{Diag}(T)|m\rangle=\delta_{n,m}\langle n|T|m\rangle$.

Proof. Since T has only a finite number of diagonals, it can be written as a finite sum of operators having only one nonzero diagonal. Using the additivity of the Dixmier trace there is no loss of generality in assuming that T has only one non-zero diagonal, namely, that it acts on $\ell^2(\mathbf{Z}^D)$ as

$$T\psi(n)=t(n)\psi(n-a), \quad \psi\in\ell^2(\mathbf{Z}^D),\ a\in\mathbf{Z}^D,$$

where t is a sequence on $\mathbf{Z}^D$ such that $|t(n)|\leqslant C(1+|n|)^{-D}$. It is thus enough to prove that $T\in\mathscr{L}^{1+}$ whatever the value of a and that, if $a\neq 0$, its Dixmier trace vanishes.

It is clear that the modulus $|T|$ of T is a diagonal operator dominated by CR where R is the multiplication operator by $(1+|n|)^{-D}$. Let us show that $R\in\mathscr{L}^{1+}$ which implies $T\in\mathscr{L}^{1+}$. Its eigenvalues are $1/j^D$ with a multiplicity $O(j^{D-1})$, therefore labeling them in decreasing order $\mu_1\geqslant\cdots\geqslant\mu_s\geqslant\cdots$ with their multiplicity, we get

$$\sup_{N>0}\frac{1}{N}\sum_{s=1}^{N}\mu_s\leqslant\sup_{N>0}\frac{1}{N}\sum_{j=1}^{O(N^{1/D})}\text{const.}\ \frac{1}{j}<\infty.$$

Thus $R\in\mathscr{L}^{1+}$.

Let us now assume $a\neq 0$. We will show that T is then unitarily equivalent to $-T$. Since the Dixmier trace is invariant by unitary transformations (Proposition 6) it will follow that $\mathrm{Tr}_{\mathrm{Dix}}(T)=0$. We remark that for any $n\in\mathbf{Z}^D$ the subspace $\mathscr{E}_n=l^2(n+a\mathbf{Z})$ is invariant under the action of T. Clearly $\mathscr{E}_n$ is isomorphic to $l^2(\mathbf{Z})$ and through this isomorphism T acts as $T_n\varphi(j)=t(n+ja)\varphi(j-1)$ on $\varphi\in l^2(\mathbf{Z})$. Let us define the unitary operator U on $l^2(\mathbf{Z})$ as the multiplication by $(-)^j$. Then one easily finds $UT_nU^{-1}=-T_n$. Lifting U to $\mathscr{E}_n$ gives a unitary denoted by U_n. Now $l^2(\mathbf{Z}^D)$ is the direct sum of the $\mathscr{E}_n$ whenever n runs in a fundamental domain of the subgroup $a\mathbf{Z}$ acting on $\mathbf{Z}^D$ by translation. Taking the direct sum of the corresponding U_n's gives a unitary $\hat{U}$ on $l^2(\mathbf{Z}^D)$ such that $\hat{U}T\hat{U}^{-1}=-T$. □

The next step in the proof of Proposition 7 is the following:

Lemma 3: Let Σ *be a subset of* $\mathbf{Z}^D$ *not including the origin and with finite density* $\mathrm{Dens}(\Sigma)$, *namely,*

$$\mathrm{Dens}(\Sigma)=\lim_{N\to\infty}\frac{1}{N^D}\sum_{n\in\Sigma,|n|\leqslant N}1,$$

where $|n|$ *is the Euclidean norm of the vector n. Then if* R_Σ *is the restriction to* Σ *of the operator of multiplication by* $1/|n|^D$ *in* $l^2(\mathbf{Z}^D)$, *we find*

$$\mathrm{Tr}_{\mathrm{Dix}}(R_\Sigma)=\frac{\omega_D}{D}\mathrm{Dens}(\Sigma),$$

where ω_D *is the area of the* $D-1$ *unit sphere of* $\mathbf{R}^D$. *In particular, for* $D=2$, *the geometrical constant in the right-hand side is* π.

Proof: The eigenvalues of R are $1/j^D$. The multiplicity $g_j(\Sigma)$ of such an eigenvalue is therefore given by the number of n's in Σ such that $|n|=j$. Let Σ_N is the subset of Σ of elements n with $|n|\leqslant N$. Since the eigenvalues of R are already labeled in decreasing order, we obtain

$$\mathrm{Tr}_{\mathrm{Dix}}(R_\Sigma)=\lim_{N\to\infty}\frac{1}{\ln|\Sigma_N|}\sum_{j=1}^{N}\frac{g_j(\Sigma)}{j^D}, \tag{48}$$

where $|\Sigma_N|$ denotes the number of points in Σ_N. Using the definition of the density of a subset, we see that, as $j\to\infty$, the multiplicity $g_j(\Sigma)$ is asymptotically given by the product of the density of Σ by the volume between the balls of radii $j-1$ and j. Namely,

$$g_j(\Sigma)\overset{j\to\infty}{\sim}\mathrm{Dens}(\Sigma)\ \ \omega_D\, j^{D-1}.$$

In much the same way, we get $|\Sigma_N|\overset{N\to\infty}{\sim}\mathrm{Dens}(\Sigma)\ \Omega_D\, N^D$ if Ω_D is the volume of the unit ball of $\mathbf{R}^D$. Taking the logarithm we are left with $D\ln N+O(1)$ in the expression of the Dixmier trace of R_Σ. Plugging all these relations in Eq. (48), we get the result. □

The last technical step in the proof of Proposition 7 is provided by the following lemma where the dimension is $D=2$.

Lemma 4: Let f be a continuous non-negative function on Ω and $a\in\mathbf{Z}^2$, $a\neq 0$. Let F^a_ω the operator on $\ell^2(\mathbf{Z}\mathbf{Z}^2)$ defined by

$$F^a_\omega\psi(n)=f(T^{-n}\omega)\left|\frac{n}{|n|}-\frac{n-a}{|n-a|}\right|^2\psi(n),\quad \psi\in\ell^2(\mathbf{Z}\mathbf{Z}^2).$$

Then $F^a_\omega\in\mathscr{L}^{1+}$ and its Dixmier trace is given by

$$\mathrm{Tr}_{\mathrm{Dix}}(F^a_\omega)=\frac{\pi}{2}|a|^2\int d\mathbf{P}(\omega)f(\omega),\quad \text{for } \mathbf{P}\text{-almost every } \omega\in\Omega.$$

Proof: As $|n|\to\infty$ the function

$$\phi(n)=\left|\frac{n}{|n|}-\frac{n-a}{|n-a|}\right|^2,$$

admits the asymptotics $\phi(n)\overset{|n|\to\infty}{\sim}|a|^2\sin^2\alpha_n/|n|^2$ modulo terms of order $1/|n|^3$, where α_n is the angle between the directions of a and n.

Let us now slice the space Ω according to the finite partition $\{\Omega_j,\delta\}$, where $\delta>0$ is small enough and j an integer such that $\Omega_{j,\delta}$ is the set of points ω for which $(j-1/2)\,\delta\leqslant f(\omega)<(j+1/2)\,\delta$. Since f is continuous with compact support, it is bounded so that only a finite number of j's are needed here. Let then $\Sigma_{j,\delta}(\omega)$ be then set of n's in $\mathbf{Z}^2$ such that $T^{-n}\omega\in\Omega_{j,\delta}$. Using Birkhoff's ergodic theorem, for $\mathbf{P}$-almost every ω, $\Sigma_{j,\delta}(\omega)$ has a finite density given by the probability $\mathbf{P}(\Omega_{j,\delta})$.

We then slice $\Sigma_{j,\delta}$ into a finite subpartition $\{\Sigma_{j,\delta},r\}$ where $\Sigma_{j,\delta},r$ corresponds to those points $n\in\Sigma_{j,\delta}$ for which $(r-1/2)\,\delta\leqslant\alpha_n<(r+1/2)\,\delta$. Thus modulo an error of order $O(\delta)$ we get $f(T^{-n}\omega)\,\phi(n)=|a|^2\,j\,\delta\,\sin^2(r\delta)(1+O(\delta))/|n|^2$ on $\Sigma_{j,\delta},r$. Moreover since this slicing concerns only a finite partition it permits to write F^a_ω as a finite sum, namely,

$$F^a_\omega=\sum_{j,r}F^a_\omega|_{\Sigma_{j,\delta,r}}.$$

It is thus sufficient to compute the Dixmier trace of the restriction to $\Sigma_{j,\delta,r}$ of F^a_ω. But up to an error of order $O(\delta)$ this restriction is nothing but $|a|^2 j\delta \sin^2(r\delta) R_{\Sigma j.\delta,r}$. Using the Lemma 3 we then get

$$\mathrm{Tr}_{\mathrm{Dix}}(F^a_\omega|_{\Sigma_{j,\delta,r}}) = \pi\ |a|^2\ j\delta\ \sin^2(r\delta)(1+O(\delta))\ \mathrm{Dens}(\Sigma_{j,\delta,r}).$$

Due to the slicing of the angles one gets $\mathrm{Dens}(\Sigma_{j,\delta,r}) = \delta/2\pi\ \mathrm{Dens}(\Sigma_{j,\delta}) = \mathbf{P}(\Omega_{j,\delta})\,\delta/2\pi$. Plugging everything together, summing up over j,r and letting δ converge to zero, the sum over r gives the averaged value of $\sin^2(\alpha)$, namely, 1/2, whereas the sum over j reconstructs the integral of f. □

Proof of Proposition 7 (end). Thanks to Lemma 2, it is enough to compute the diagonal elements of $|\hat{\delta}A_\omega|^2$ because $A \in \mathscr{A}_0$ so that the number of nonzero diagonals is finite. The diagonal elements are

$$\langle n||\hat{\delta}A_\omega|^2|n\rangle = \sum_{a\in\mathbf{Z}^2} |A(T^{-n}\omega,a)|^2 \left|\frac{n}{|n|} - \frac{n-a}{|n-a|}\right|^2.$$

The number of terms in this sum is finite. Using Lemma 4 we find

$$\mathrm{Tr}_{\mathrm{Dix}}(|\hat{\delta}A_\omega|^2) = \frac{\pi}{2} \sum_{a\in\mathbf{Z}^2} |a|^2 \int_\Omega d\mathbf{P}(\omega)|A(\omega,a)|^2,$$

for $\mathbf{P}$-almost all ω's. On the other hand, the definition of the differential on $\mathscr{A}$ gives [see Eq. (33)] $\nabla A(\omega,a) = i\mathbf{a}A(\omega,a)$. In particular

$$|\nabla A|^2(\omega,0) = \sum_{a\in\mathbf{Z}^2} |a|^2 |A(\omega,a)|^2.$$

To get the trace per unit volume, we just have to integrate both sides of this equation giving the Connes formula.

Since the left-hand side of the Connes formula is dominated by the Sobolev norm of A, one can extend this formula to $A \in \mathscr{S}$. In particular the finiteness of the right-hand side implies that $\hat{\delta}A_\omega \in \mathscr{L}^{2+}$ $\mathbf{P}$-almost surely. □

E. Noncommutative Chern character

We denote by $\mathscr{P}(\mathscr{A})$ the set of orthogonal projections in the C^*-algebra $\mathscr{A}$, namely, the set of elements P in $\mathscr{A}$ such that $P = P^2 = P^*$. If in addition P is differentiable, we define its Chern character as

$$\mathrm{Ch}(P) = 2\imath\pi\mathscr{T}(P[\partial_1 P,\ \partial_2 P]). \tag{49}$$

If we work on a lattice, $\mathscr{T}$ is normalized such that $\mathscr{T}(\mathbf{1}) = 1$. In the continuum case, we will normalize it in reference to the projections onto Landau levels, namely the eigenvalues of the Landau Hamiltonian. (5). An elementary calculation gives for the lowest Landau level projection the following integral kernel

$$\Pi_0(\mathbf{x},\mathbf{y}) = \frac{q\mathscr{B}}{h}\, e^{-(q\mathscr{B}/4\hbar)(\mathbf{x}-\mathbf{y})^2 - \imath(q\mathscr{B}/2\hbar)\mathbf{x}\wedge\mathbf{y}}. \tag{50}$$

In particular, Π_0 defines an element of $\mathcal{A}=C^*(\Omega,\mathbf{R}^2,\mathcal{B})$ for any choice of the hull Ω. In much the same way, we denote by Π_n the projection on the n^{th} Landau level ($n=0,1,2,...$). We then deduce the following results.

Lemma 5: The trace and the Chern character of the Landau levels are given by

$$\mathcal{T}(\Pi_n)=\frac{q\mathcal{B}}{h}, \quad \mathrm{Ch}(\Pi_n)=-1. \tag{51}$$

Proof. Let us prove this result for Π_0 first. Its trace per unit volume is given by the space average of $\Pi_0(\mathbf{x},\mathbf{x})=q\mathcal{B}/h$. This gives the first formula in (51). To compute its Chern character, we remark that $\nabla\Pi_0(\mathbf{x},\mathbf{y})=-\imath(\mathbf{x}-\mathbf{y})\Pi_0(\mathbf{x},\mathbf{y})$. Introducing the complex variables $x=x_1+\imath x_2$ and $y=y_1+\imath y_2$, we get

$$\mathrm{Ch}(\Pi_0)=\pi\left(\frac{q\mathcal{B}}{h}\right)^3\int_{\mathbf{C}\times\mathbf{C}}d^2xd^2ye^{-(q\mathcal{B}/2\hbar)(|x|^2+|y|^2-x\bar{y})}(x\bar{y}-y\bar{x}).$$

To compute this integral, we develop the exponential in powers of $x\bar{y}$ and notice that all contributions vanish except the one involving the term $|x|^2|y|^2$. The corresponding integral has separated variables and can be computed explicitly. This gives -1.

For the other Landau levels, we remark that $\Pi_n\sim\Pi_0$ in the sense of von Neumann equivalence (see below). More concretely, it is enough to exhibit an element $U_n\in\mathcal{A}$ such that $\Pi_n=U_n^*U_n$ and $\Pi_0=U_nU_n^*$. This implies that the traces are identical and we will show their Chern characters to be identical as well (see Lemma 11 below). To construct U_n, we introduce the annihilation operator

$$a=(P_1-qA_1+\imath P_2-qA_2)/\sqrt{2\hbar q\mathcal{B}}.$$

Then $[a,\ a^*]=\mathbf{1}$. Thus aa^* is bounded below by 1 and is invertible. We set $u=(aa^*)^{-1/2}a$ and $U_n=\Pi_0u^n$. It is easy to check that $uu^*=\mathbf{1}$ implying that $\Pi_0=U_nU_n^*$. On the other hand, $U_n^*U_n=a^{*n}\Pi_na^n/n!$. But it is a standard result that we obtain the nth Landau level by applying the creation operator n times to the ground state, namely, $|n>=(1/\sqrt{n!})a^{*n}|0\rangle$. In particular $a^{*n}\Pi_na^n/n!=\Pi_n$.

It remains to show that $U_n\in\mathcal{A}$. A straightforward calculation gives $U_n=(n!)^{-1/2}\Pi_0a^n$. Now using the explicit form of the matrix elements (50) and of a, we get the matrix elements of U_n in the form of a polynomial in x and y times $\Pi_0(\mathbf{x},\mathbf{y})$, showing that $U_n(\mathbf{0},\mathbf{x})$ is absolutely summable in $\mathbf{x}$. This is enough to show that it belongs to $\mathcal{A}$ (see Sec. III F). □

Our next step will be the von Neumann equivalence. Namely, if $P,Q\in\mathcal{P}(\mathcal{A})$, then $P\sim Q$ if there is $U\in\mathcal{A}$ such that $P=U^*U, Q=UU^*$. In particular if P is trace class, it follows that $\mathcal{T}(P)=\mathcal{T}(Q)$ then. The following results can be found in Ref. 77.

Lemma 6: Let $P,Q\in\mathcal{P}(\mathcal{A})$ be such that $\|P-Q\|<1$. Then $P\sim Q$.

Lemma 7: If $\mathcal{A}$ is separable, then the set of equivalence classes of projections in $\mathcal{A}$ is at most countable.

Lemma 8: Let $P,Q\in\mathcal{P}(\mathcal{A})$ be two mutually orthogonal projections (namely, $PQ=QP=0$). Then the equivalence class of their sum $P\oplus Q$ depends only upon the equivalence classes of P and Q. This defines a commutative and associative composition law on the set of equivalence classes, which we denote by $[P\oplus Q]=[P]+[Q]$.

See Refs. 79 and 61, Lemma 4.2.3 for a proof.

Remark that we can add only mutually orthogonal projections in this way, because the sum of two projections is not a projection in general. So giving any pair P,Q of projections it is not always possible to find equivalent projections $P'\sim P$ and $Q'\sim Q$ such that P' and Q' be orthogonal. In other words the sum is not everywhere defined in $\mathcal{P}(\mathcal{A})$. In order to deal with this problem, we replace the algebra $\mathcal{A}$ by its *stabilization*, namely, the tensor product $\mathcal{A}\otimes\mathcal{K}$ with the algebra of compact operators on a separable Hilbert space, which is nothing but the smallest C^*-algebra containing all finite dimensional matrices. Then it is possible to show that one can always choose pairs of projection as orthogonal up to equivalence. By the Grothendieck method one builds a group out of the equivalence classes of projections of $\mathcal{A}\otimes\mathcal{K}$. This group is denoted $K_0(\mathcal{A})$ (see Refs. 78 and 61, Theorem 10.

Lemma 9: For any $P\in\mathcal{P}(\mathcal{A})$ there is a differentiable projection $P_0\in\mathcal{P}(\mathcal{A})$ such that $P\sim P_0$.

This last result is a consequence of the fact that the set of differential elements in $\mathcal{A}$ is norm dense.

Lemma 10: For any pair $P,Q\in\mathcal{P}(\mathcal{A})$ of equivalent projections and any ε small enough, there are differentiable projections $P_\varepsilon,Q_\varepsilon\in\mathcal{P}(\mathcal{A})$ and a differentiable element U_ε such that $\|P_\varepsilon-P\|\leqslant\varepsilon$, $\|Q_\varepsilon-Q\|\leqslant\varepsilon$ and $P_\varepsilon=U_\varepsilon U_\varepsilon^*$ whereas $Q_\varepsilon=U_\varepsilon^* U_\varepsilon$.

The proof is straightforward: it follows from the density of $\mathcal{C}^1(\mathcal{A})$ and from the proof of Lemma 6. We will say that P and Q are smoothly equivalent whenever the element U which connects them can be chosen differentiable.

Lemma 11: For any pair $P,Q\in\mathcal{P}(\mathcal{A})$ of smoothly equivalent projections $\mathrm{Ch}(P)=\mathrm{Ch}(Q)$.

The proof of this result is purely combinatorial provided we use the cyclicity of the trace $\mathcal{T}$. We will omit it here.[4]

Lemma 12: For any pair $P,Q\in\mathcal{P}(\mathcal{A})$ of mutually orthogonal smooth projections we have $\mathrm{Ch}(P\oplus Q)=\mathrm{Ch}(P)+\mathrm{Ch}(Q)$.

The proof of this result is standard and can be found in Ref. 4 for instance. To summarize this set of results we have

Theorem 8: *The Chern character Ch defines a group homomorphism from $K_0(\mathcal{A})$ into a countable subgroup of the real line.*

It remains to show that the image of this map is the set of integers. This will be done in Secs. IV F and IV G.

F. Connes formulas

In order to compute eventually the Chern character of a projection, we need an intermediate tool, namely, a cyclic cocycle. This is actually the heart of Connes work on the noncommutative extension of cohomology. Here, we shall actually only need a 2-cocycle τ_2, which is a trilinear form on the algebra $\mathcal{A}_0$ defined by

$$\mathcal{T}_2(A_0,A_1,A_2)=2\pi i\mathcal{T}(A_0\partial_1 A_1\partial_2 A_2-A_0\partial_2 A_1\partial_1 A_2). \tag{52}$$

By the Schwarz inequality for the trace τ, we see that τ_2 can be extended to the noncommutative Sobolev space $\mathcal{S}$, which is linear subspace of the von Neumann algebra $\mathcal{W}$. The proof of the following lemma is algebraic and standard by now; it can be found in Ref. 4.

Lemma 13: $\mathcal{T}_2$ is a 2-cocycle, i.e., it satisfies the following algebraic properties:

(i) $\mathscr{T}_2$ *is cyclic:* $\mathscr{T}_2(A_0,A_1,A_2)=\mathscr{T}_2(A_2,A_0,A_1)$,
(ii) $\mathscr{T}_2$ *is closed under Hochschild's boundary operator, that is: b:*

$$(b\mathscr{T}_2)(A_0,A_1,A_2,A_3)\equiv\mathscr{T}_2(A_0A_1,A_2,A_3)-\mathscr{T}_2(A_0,A_1A_2,A_3)+\mathscr{T}_2(A_0,A_1,A_2A_3)$$
$$-\mathscr{T}_2(A_3A_0,A_1,A_2)=0.$$

We will now give a formula which permits to compute the cocycle τ_2 by means of the physical representations. For this purpose, we present the formalism introduced by Connes[4] (we already gave some indications in Sec. III B). A graded Fredholm module is defined as follows. Let $\mathscr{H}_+$ and $\mathscr{H}_-$ be two separable Hilbert spaces; their direct sum $\hat{\mathscr{H}}=\mathscr{H}_+\oplus\mathscr{H}_-$ becomes a graded Hilbert space through the graduation operator $\hat{G}$ equal to ± 1 on $\mathscr{H}_\pm$. A representation $\hat{\pi}:\mathscr{D}\to\mathscr{L}(\hat{\mathscr{H}})$ of an algebra $\mathscr{D}$ is said to be trivially graded, if $[\hat{\pi}(A),\hat{G}]=0$ for all $A\in\mathscr{D}$. One says that $\hat{\pi}(A)$ is of degree 0; operators on $\hat{\mathscr{H}}$ which anticommute with the graduation operator $\hat{G}$ are said to be of degree 1. Any operator on $\hat{\mathscr{H}}$ can be uniquely decomposed into the sum of an operator of degree 0 with an operator of degree 1.

Definition 6: A Fredholm module is a family $(\mathscr{D},\hat{\mathscr{H}},\hat{\pi},F)$ *where* $\mathscr{D}$ *is an algebra with a trivially graded representation* $\hat{\pi}$ *on a graded, separable Hilbert space* $\hat{\mathscr{H}}$ *and where* $F\in\mathscr{L}(\hat{\mathscr{H}})$ *is a self-adjoint operator such that the following three conditions are satisfied:*

$$(i)\ F\hat{G}=-\hat{G}F,\quad (ii)\ F^2=1,\quad (iii)\ [\hat{\pi}(A),F]\in\mathscr{K}\forall A\in\mathscr{D}.$$

Here $\mathscr{K}$ *is the ideal of compact operators on* $\hat{\mathscr{H}}$. *An element* $A\in\mathscr{D}$ *is called p-summable (resp. p+-summable) whenever* $[\hat{\pi}(A),F]\in\mathscr{L}^p(\hat{\mathscr{H}})$ $(resp.\in\mathscr{L}^{p+}(\hat{\mathscr{H}}))$.

The graded commutator of two graded operators $T,T'\in\mathscr{L}(\hat{\mathscr{H}})$ *is defined by*

$$[T,\ T']_S=TT'-(-1)^{d^o(T)d^o(T')}T'T.$$

This commutator extends to the whole algebra $\mathscr{L}(\hat{\mathscr{H}})$ by bilinearity. The noncommutative differential is given by

$$dT=[F,\ T]_S,\quad T\in\mathscr{L}(\hat{\mathscr{H}}).$$

One can check that it obeys the graded Leibniz rule $d(TT')=dT\ T'+(-)^{d^o(T)}T\ dT'$ and that $d^2=0$. Finally, the graded trace or the supertrace is defined by

$$\mathrm{Tr}_S(T)=\tfrac{1}{2}\mathrm{Tr}(\hat{G}FdT),$$

whenever the right-hand side is well defined. Here, Tr is the usual trace in $\mathscr{L}(\hat{\mathscr{H}})$. Remark that Tr_S is linear and satisfies:

$$\mathrm{Tr}_S(TT')=(-1)^{d^o(T')d^o(T)}\mathrm{Tr}_S(T'T).$$

Moreover, if $d^o(T)=1$, then $\mathrm{Tr}_S(T)=0$.

Now we shall consider the concrete family of Fredholm modules which will be of interest to us. The algebra will be $\mathscr{A}_0$. The Hilbert space is $\hat{\mathscr{H}}=\mathscr{H}^+\oplus\mathscr{H}^-=L^2(G)\oplus L^2(G)$ (where $G=\mathbf{R}^2$ or $G=\mathbf{Z}^2$) and for $A\in\mathscr{A}_0$ the representation is given by

$$\hat{\pi}_\omega(A)=\hat{A}_\omega=\begin{pmatrix}A_\omega & 0\\ 0 & A_\omega\end{pmatrix},\qquad A_\omega=\pi_\omega(A),$$

where π_ω is the family of representations defined in (31). Next

$$F=\begin{pmatrix} 0 & F^{+-} \\ F^{-+} & 0 \end{pmatrix}=\begin{pmatrix} 0 & u \\ u^* & 0 \end{pmatrix}, \qquad u=\frac{X_1+\imath X_2}{|X_1+\imath X_2|},$$

here X_1, X_2 are the two components of the position operator on $L^2(G)$.

The following result holds for $G=\mathbf{Z}^2$ or $\mathbf{R}^2$. However, we will prove it only in the discrete case and leave the continuous case for a forthcoming work.[64]

Theorem 9 (First Connes formula): *The Fredholm module $(\mathcal{A}_0, \hat{\mathcal{H}}, \hat{\pi}_\omega, F)$ defined above is 2+-summable for* **P***-almost all ω's. Moreover for every $A \in \mathcal{A}_0$, the following formula holds:*

$$\mathcal{T}(|\nabla A|^2)=\frac{1}{\pi}\,\mathrm{Tr}_{\mathrm{Dix}}(|dA_\omega|^2), \quad \textit{for } \mathbf{P}\textit{-almost all } \omega. \tag{53}$$

If $\mathcal{S}$ denotes the Sobolev space associated to $\mathcal{T}$, this formula can be continued to elements $A \in \mathcal{S}$. In particular, if $A \in \mathcal{S}$, then $dA_\omega \in \mathcal{L}^{2+}$ for **P***-almost all ω.*

This result is an elementary extension of Proposition 7. Its proof is left to the reader.

The following formula links the cocycle τ_2 defined in (52) to the previous Fredholm modules $(\mathcal{A}_0, \hat{\mathcal{H}}, \hat{\pi}_\omega, F)$; it can already be found in Ref. 4. This is a result specific to the algebra $\mathcal{A}_0$ and depends upon the dimension $D=2$ of the physical space (or of the Brillouin zone).

Theorem 10 (Second Connes formula): *For $A_0, A_1, A_2 \in \mathcal{A}_0$, we have the following formula:*

$$\int_\Omega d\mathbf{P}(\omega)\,\mathrm{Tr}_S(\hat{A}_{0,\omega} d\hat{A}_{1,\omega} d\hat{A}_{2,\omega})=\mathcal{T}_2(A_0, A_1, A_2). \tag{54}$$

Proof: We remark that the left-hand side is well defined thanks to Theorem 9. For indeed, if $A \in \mathcal{A}_0$ then $dA_\omega \in \mathcal{L}^{2+} \subset \mathcal{L}^3$ **P**-almost surely. The trace of an integral operator on $L^2(G)$ with continuous and compactly supported integral kernel is given by the integral of its diagonal. We can therefore evaluate the left-hand side:

$$\begin{aligned} &\int_\Omega d\mathbf{P}(\omega)\,\mathrm{Tr}_S(\hat{A}_{0,\omega} d\hat{A}_{1,\omega} d\hat{A}_{2,\omega}) \\ &= \int_\Omega d\mathbf{P}(\omega) \int_{G^3} d^2x_0 d^2x_1 d^2x_2 \left[-\left(1-\frac{\bar{x}_0}{|x_0|}\frac{x_1}{|x_1|}\right)\left(1-\frac{\bar{x}_1}{|x_1|}\frac{x_2}{|x_2|}\right)\left(1-\frac{\bar{x}_2}{|x_2|}\frac{x_0}{|x_0|}\right)\right] \\ &\quad \times e^{\imath\lambda(x_0 \wedge x_1 + x_1 \wedge x_2 + x_2 \wedge x_0)} A_0(T^{-x_0}\omega, x_1-x_0) A_1(T^{-x_1}\omega, x_2-x_1) A_2(T^{-x_2}\omega, x_0-x_2), \end{aligned} \tag{55}$$

where $\lambda = qB/2\hbar$. The main ingredient of the proof is now the following lemma for which there are two different proofs in Refs. 4 and 17. We shall follow Ref. 17, but present a discrete version of the proof; for the continuous case $G=\mathbf{R}^2$ we refer to 17.

Lemma 14: Let $a,b \in G$ which we write as $a=a_1+\imath a_2$, $b=b_1+\imath b_2$. Then we have

$$-2\pi\imath a \wedge b=\sum_{s \in G}\left(1-\frac{\bar{s}}{|s|}\frac{s-a}{|s-a|}\right)\left(1-\frac{\overline{s-a}}{|s-a|}\frac{s-b}{|s-b|}\right)\left(1-\frac{\overline{s-b}}{|s-b|}\frac{s}{|s|}\right).$$

Proof: In the discrete case $G=\mathbf{Z}^2$, the integral appearing in the lemma is in fact a sum, let us denote it by $C(a,b)$. Then $C(a,b)=-C(b,a)$ and $\overline{C(a,b)}=-C(a,b)$. Hence $C(a,b)$ is purely imaginary. Now we define:

$$e(s,t)=\left(\frac{s}{|s|}\,\frac{\bar{t}}{|t|}-\frac{t}{|t|}\,\frac{\bar{s}}{|s|}\right)=-e(t,s)=-\overline{e(s,t)}.$$

A direct calculation leads to

$$C(a,b)=-\sum_{s\in\mathbf{Z}^2}\left(e(s-a,s-b)+e(s-b,s)+e(s,s-a)\right).$$

We introduce $C_N(a,b)$ as the same sum in which s is restricted to be smaller or equal to $N \in \mathbf{N}$. The finite difference operators are defined with help of the translation operators T_1, T_2 on functions on $\mathbf{Z}^2$ as

$$\Delta_j=T_j-1, \quad \tilde{\Delta}_j=1-T_j^{-1}, \quad j=1,2.$$

We consider the finite differences of of $C_N(a,b)$:

$$(\Delta_{a_1}\Delta_{b_2}-\Delta_{b_1}\Delta_{a_2})C_N(a,b)=-\sum_{|s|\leqslant N}\tilde{\Delta}_{s_1}\frac{s-a}{|s-a|}\,\tilde{\Delta}_{s_2}\frac{\overline{s-b}}{|s-b|}-\tilde{\Delta}_{s_1}\frac{s-b}{|s-b|}\,\tilde{\Delta}_{s_2}\frac{\overline{s-a}}{|s-a|}+(1\leftrightarrow 2).$$

A discrete analog of Stokes' theorem allows us to transform the sum over the square into a sum over the border of the square. As $N\to\infty$, this sum converges to the Riemann integral $\int_0^{2\pi} e^{\imath\phi}\,de^{-\imath\phi}$. The term $(1\leftrightarrow 2)$ gives the same contribution and we obtain

$$(\Delta_{a_1}\Delta_{b_2}-\Delta_{b_1}\Delta_{a_2})C(a,b)=-4\pi\imath.$$

As this is true for every $a,b\in\mathbf{Z}^2$, $C(a,b)$ is of the form

$$C(a,b)=\alpha+\beta(a,b)-2\pi\imath a\wedge b,$$

where α is a constant and β is linear in the vector (a,b). As $C(0,0)=0$, we have $\alpha=0$. Because $C(a,b)$ and $a\wedge b$ are both odd under permutation of a and b, β must be odd as well so that it vanishes. □

Proof of Theorem 10 (end): We set $s=x_0, s_0=x_0-x_1, s_1=x_1-x_2$ in Eq. (55) and use the invariance of the measure $\mathbf{P}$ in order to replace $T^{-s}\omega$ by ω. Applying Lemma 14 we get

$$\int_\Omega d\mathbf{P}(\omega)\mathrm{Tr}_S(\hat{A}_{0_\omega}d\hat{A}_{1_\omega}d\hat{A}_{2_\omega})=-\int_\Omega d\mathbf{P}(\omega)\int d^2s_0d^2s_1e^{\imath\lambda s_0\wedge s_1}2\pi\imath s_0\wedge s_1A_0(\omega,-s_0)A_1$$
$$\times(T^{s_0}\omega,-s_1)A_2(T^{s_0+s_1}\omega,s_0+s_1).$$

The right-hand side is precisely the formula for $\mathcal{T}_2(A_0,A_1,A_2)$. □

G. Chern character and Fredholm index

The main interest of Connes' theory of noncommutative Fredholm modules[4] is the following in our context: for a given 3-summable projection P in an algebra $\mathcal{A}$, the expression $\sigma(P,P,P)=\frac{1}{2}\mathrm{Tr}_S(\hat{P}d\hat{P}d\hat{P})$ can be related to the index of a Fredholm operator. In order to make this article self-contained, we will reproduce here the main steps relevant for us.

First we need the following formula due to Fedosov.[80] A proof can be found in the appendix of Ref. 4.

Proposition 8 (Fedosov's formula): Let F be a bounded operators on a Hilbert space $\mathscr{H}$. We suppose that $(1-F^*F)\in\mathscr{L}^p(\mathscr{H})$ and $(1-FF^*)\in\mathscr{L}^p(\mathscr{H})$ for some $p\in[1,\infty)$. Then F is a Fredholm operator and for every integer $n\geq p$ its index satisfies

$$\mathrm{Ind}(F)=\mathrm{Tr}((1-F^*F)^n)-\mathrm{Tr}((1-FF^*)^n).$$

Proposition 9: Let $(\mathscr{D},\hat{\mathscr{H}},\hat{\pi},F)$ be a Fredholm module and $P\in\mathscr{D}$ be a 3-summable projection. Then $F_P^{+-}=PF^{+-}|_{P\mathscr{H}^-}$ is a Fredholm operator and

$$\mathrm{Ind}(F_P^{+-})=\mathrm{Tr}_S(\hat{P}d\hat{P}d\hat{P}),\quad \hat{P}=\begin{pmatrix}\pi(P) & 0\\ 0 & \pi(P)\end{pmatrix}.$$

Proof: Suppose $\mathscr{H}^+=\mathscr{H}^-$ for simplicity. Then

$$-\hat{P}[F,\hat{P}]^2\hat{P}=\hat{P}-\hat{P}F\hat{P}F\hat{P}=\begin{pmatrix}(1-F_P^{+-}F_P^{-+})|_{P\mathscr{H}^+} & 0\\ 0 & (1-F_P^{-+}F_P^{+-})|_{P\mathscr{H}^-}\end{pmatrix}.$$

By hypothesis $[F,\hat{P}]\in\mathscr{L}^3(\hat{\mathscr{H}})$. Hölder's inequalities imply $(1-F_P^{+-}F_P^{-+})|_{P\mathscr{H}^+}\in\mathscr{L}^2(\mathscr{H}^+)$ and $(1-F_P^{-+}F_P^{+-})|_{P\mathscr{H}^-}\in\mathscr{L}^2(\mathscr{H}^-)$. By Fedosov's formula we get

$$\mathrm{Ind}(F_P^{+-})=\mathrm{Tr}_{P\mathscr{H}^-}((1-F_P^{-+}F_P^{+-})^2)-\mathrm{Tr}_{P\mathscr{H}^+}((1-F_P^{+-}F_P^{-+})^2)=-\mathrm{Tr}_{\hat{\mathscr{H}}}(\hat{G}(\hat{P}-\hat{P}F\hat{P}F\hat{P})^2).$$

We can check that this is equal to $\mathrm{Tr}_S(\hat{P}d\hat{P}d\hat{P})$ by using the following algebraic identities

$$[F,\hat{P}]=\hat{P}[F,\hat{P}]+[F,\hat{P}]\hat{P},\quad \hat{P}[F,\hat{P}]^2=[F,\hat{P}]^2\hat{P},$$

$$F[F,\hat{P}]^{2n+1}=-[F,\hat{P}]^{2n+1}F,\quad F\hat{G}=-\hat{G}F.$$

The proof can easily be completed. □

We shall now extend this result to stochastic operators.

Theorem 11 (Ref. 16): *Let $P\in\mathscr{W}$ be a projection belonging to the noncommutative Sobolev space $\mathscr{S}$. Then for* **P***-almost every $\omega\in\Omega$, P is 2+-summable and*

$$\mathbf{Ch}(P)=\mathrm{Ind}(P_\omega u|_{P_\omega(\mathscr{H}^-)}),$$

where $P_\omega=\pi_\omega(P)$ and $u=X/|X|$. In particular, $\mathbf{Ch}(P)$ *is an integer.*

Proof: Using Theorem 9, Theorem 10, and Proposition 9, it just remains to show that the index is **P**-almost surely independent of ω. Using the ergodicity of **P,** it is enough to show that the index is translation invariant. For indeed, translating P_ω by $a\in G$ just changes ω into $T^{-a}\omega$. On the other hand, translating u by a changes it into $u+O(1/|X|)$. Thus $P_\omega u|_{P_\omega(\mathscr{H}^-)}$ is changed into $P_{T^{-a}\omega}u|_{P_{T^{-a}\omega}(\mathscr{H}^-)}$ modulo a compact operator. Since a compact perturbation does not change the index of a Fredholm operator, the result is achieved. □

Remark. The above theorem is true independently of the choice of the probability measure **P**. However, changing **P** is equivalent to changing the disorder. Therefore we cannot expect the Sobolev condition to hold independently of **P**. In particular, if H is a given bounded self-adjoint operator, the spectrum of its representative $\pi_\omega(H)$ is **P**-almost surely constant,[31] but changing **P** may change it. Therefore, **P** has a physical content. We will see in the Sec. VI C some of the consequences of changing the disorder. □

H. Quantization, Fredholm, and relative index

In this section, we will discuss the links between the Laughlin argument as presented in Sec. II E and our approach. The essential ingredient for that will be the relative index of two projections as defined by Avron, Seiler, and Simon.[17] It turns out that the singular gauge transformation of Laughlin corresponds to the unitary operator $u=X/|X|$. Thus the charge transported after changing the flux by one quantum is then exactly given by the index we computed in Theorem 11. The main improvement upon Laughlin's argument is that we control completely the effect of the disorder now, since this index is a topological invariant. Let us remark that the topology we are talking about is the one of the Brillouin zone and not of the sample as can be erroneously derived from a superficial understanding of Laughlin's argument.

More precisely, following Sec. II E, let the varying flux be $\phi(t)=\hbar t/e\tau$ with τ so large as to produce an adiabatic change. Then at time $t=\tau$ the new Hamiltonian in (8) is given by

$$H_{\mathcal{A}}(\tau)=uH_{\mathcal{A}}(0)u^*, \quad u=\frac{X}{|X|}, \tag{56}$$

namely, the phase of the wave function changes by $e^{i\theta}$ where θ is the polar angle of the position x. Formally, one gets an equivalent formula at each intermediate times, but the domains change with time. The Eq. (56) implies for the corresponding Fermi projections:

$$P_F(\tau)=uP_F(0)u^*, \quad P_F(t)=\chi_{\leq E_F}(H_{\mathcal{A}}(t)). \tag{57}$$

The charge transported to infinity after this adiabatic change is the number of states in $P_F(\tau)$ that are not in $P_F(0)$. Since both projections are infinite dimensional, we must be careful in computing this number. It is the purpose of the following definition[17] to take care of this difficulty. Namely, given two projections P and Q on a Hilbert space we set

$$\mathrm{Index}(P,Q)=\dim(\mathrm{Ker}(P-Q-1))-\dim(\mathrm{Ker}(Q-P-1)),$$

whenever the right side is well defined. Then (see Ref. 17 for a proof)

Proposition 10: Let P and Q be two projections on Hilbert space. If $(P-Q)\in\mathscr{L}^p$, *then for every integer m such that* $2m+1\geq p$, *the relative index can be calculated as*

$$\mathrm{Index}(P,Q)=\mathrm{Tr}((P-Q)^{2m+1}).$$

We then apply this formalism to the $2+$-summable Fermi projections $P_{F,\omega}$ and $uP_{F,\omega}u^*$. It is an immediate consequence of the proof of Proposition 9 that

$$\mathrm{Index}(P_{F,\omega}uP_{F,\omega}u^*)=-\mathrm{Tr}_S(\hat{P}_{F,\omega}d\hat{P}_{F,\omega}d\hat{P}_{F,\omega}), \tag{58}$$

which is **P**-almost surely equal to $-\mathbf{Ch}(P_F)$. The suitable $p=2m+1$ is $p=3$ here. We remark that it is the smallest possible one giving a nonzero relative index because, if $P_F-uP_Fu^*$ were trace class, the relative index would vanish. Hence the Chern character of the Fermi projection can be identified up to sign, with the charged transported at infinity during the adiabatic flux change. Let us then call this index the *charge deficiency index.*

Owing to the stability of the Fredholm index, if one pierces the flux tube at some other place than the origin and then uses another unitary operator than $u=e^{i\theta}$, the Fredholm index will not change. Moreover, adding some disorder potential to the Landau Hamiltonian will not change the index as well as long as the Fermi projection belongs to the Sobolev space (we will discuss this condition more precisely in the next section).

V. LOCALIZATION AND NONCOMMUTATIVE SOBOLEV SPACE

In this chapter, we relate the Sobolev condition on Fermi projections to the Anderson localization. We will give several mathematical tools to describe rigorously localization in terms of our formalism. The main results are Theorems 13 and 14 below. As a consequence, the remaining results in the main Theorem 1 follow.

In Sec. V A we review the well-known Anderson–Pastur criterion for localization. Then we give a new definition of the localization length and discuss its relation to other notions of localization. Notice, however, that a good part of Sec. V B has already been published in Ref. 80. Our noncommutative localization length allows us to formulate a mathematically precise sufficient condition for the existence of the plateaux. All the tools we introduce fit well within the noncommutative framework developed so far. For technical simplicity, however, we will restrict ourselves to the lattice case, namely, for $G=\mathbf{Z}^D$ (tight binding representation) in this section. Even though we believe that most of these results hold for the continuum case as well, the proofs are more difficult and will be postponed to a future work.[64]

Our first definition of the localization length starts by demanding that the mean square distance that a particle moves from a given point in an infinite amount of time be finite. We will give a precise definition of the word "mean" we use here. In Theorem 13, we show that provided this condition holds, the localization length can be defined as a $L^2(\mathbf{R},d\mathcal{N})$-function of the energy, where $\mathcal{N}$ is the DOS of the Hamiltonian under study. Furthermore we give another definition of the localization length, based upon the Sobolev norm of the eigenprojections, and show in Proposition 13 that it is equivalent to the first one. Then Theorem 13 shows that if this localization length is finite in some energy interval, the spectrum of the Hamiltonian in this interval is pure point for almost all disorder configurations. Theorem 14 shows that under the same condition, the spectral projection P_E on energies lower than or equal to E is a continuous function of E with respect to the Sobolev norm. This technical result, together with the results of the previous chapter, implies that the Chern number of P_E is constant on that interval. This is the reason why we get the plateaux of the Hall conductance.

Let us remark that our localization length can be computed as a disorder average of a product between two Green functions, or equivalently by means of a current–current correlation. In particular, let m be the positive measure on $\mathbf{R}^2$ defined by[81]

$$\int_{\mathbf{R}^2} dm(E,E')f(E)g(E')=\mathcal{T}(\nabla Hf(H)\nabla Hg(H)), \quad f,g\in\mathcal{C}_0(\mathbf{R}).$$

Using spectral theory in $L^2(\mathcal{A},\mathcal{T})$, one can indeed show that such a measure exists and that it can be calculated using Green's functions. Then the localization condition that we define below is given by

$$l^2(\Delta)=2\int_{\Delta\times\mathbf{R}}\frac{dm(E,E')}{(E-E')^2},$$

whenever it exists. We will not develop this point of view here, but we emphasize that our approach is equivalent to the one used in solid state physics.

A. The Anderson–Pastur localization criterion

Most results in this section are due to Pastur.[82,33] The underlying physical idea can be traced back to Anderson.[28] Since this theory holds in any dimension, we will assume that the lattice in space is $\mathbf{Z}^D$. Let then Ω be a compact space endowed with an action of $\mathbf{Z}^D$ by homeomorphisms. The magnetic field $\mathcal{B}$ in D-dimension is an antisymmetric bilinear form on $\mathbf{Z}^D$ written as $\mathcal{B}a\wedge b$. We define the C^*-algebra $\mathcal{A}=C^*(\Omega\times\mathbf{Z}^D,\mathcal{B})$ as in Sec. III F. In this discrete case, this algebra has a unit. Given a $\mathbf{Z}^D$-invariant ergodic probability measure $\mathbf{P}$ on Ω, we get a trace $\mathcal{T}$ on

$\mathcal{A}$, which is actually normalized. We will denote by $\mathcal{W}$ the von Neumann algebra $L^\infty(\mathcal{A},\mathcal{T})$ of the corresponding GNS representation. Recall that the representations π_ω extend to $\mathcal{W}$ and give random operators. Then, to avoid inessential difficulties, the Hamiltonian H will be a self-adjoint element of $\mathcal{A}$ in this section. We will denote by $\sigma(\omega)$ the spectrum of the operator $\pi_\omega(H)=H_\omega$. As a preliminary, let us recall the

Lemma 15 (Wiener criterion): Let μ be a finite complex measure on the real line, i.e., μ is a linear combination of four finite positive measures $\mu=\mu_1-\mu_2+\imath\mu_3-\imath\mu_4$. Let $F_\mu(t)=\int e^{itx}\,d\mu(t)$ be its Fourier transform, then

$$\lim_{T\to\infty}\int_0^T \frac{dt}{T}\,|F_\mu(t)|^2=\sum_{E\in\mathbf{R}}|\mu\{E\}|^2.$$

A proof for a positive measure can be found in Ref. 76; it can be completed to a proof of Lemma 15 without any difficulties.

The next result is given by

Proposition 11: Let $P\in\mathcal{W}$, the von Neumann algebra $L^\infty(\mathcal{A},\mathcal{T})$ generated by $\mathcal{A}$ in the GNS representation of the trace per unit volume $\mathcal{T}$ associated to the probability measure **P**. *Then* **P**-*almost surely, $\pi_\omega(P)$ is a projection on $\ell^2(\mathbf{Z}^D)$. Moreover its dimension* $\dim(\pi_\omega(P))=\mathrm{Tr}(\pi_\omega(P))$ *is either zero or infinity* **P**-*almost surely.*

Proof: Because of the covariance of P, $\dim(\pi_\omega(P))$ is translation invariant. Let us define $\Pi=\int d\mathbf{P}(\omega)\pi_\omega(P)$ as a weak integral acting on $\ell^2(\mathbf{Z}^D)$. By construction, Π commutes to the translation group. Since the measure **P** is invariant and ergodic, using the monotone convergence theorem, we get for **P**-almost all $\omega\in\Omega$:

$$\dim(\pi_\omega(P))=\int d\mathbf{P}(\omega)\mathrm{Tr}(\pi_\omega(P))=\mathrm{Tr}(\Pi)=\sum_{x\in\mathbf{Z}^D}\langle x|\Pi|x\rangle=\sum_{x\in\mathbf{Z}^D}\langle 0|\Pi|0\rangle,$$

and this last expression is either zero or infinity. □

Let now $\Delta\subset\mathbf{R}$ be a Borel subset of the spectral axis. Let then $P(\Delta)$ be the corresponding spectral projection of H. Let $A_n(\Delta,\omega)$ be the time-averaged probability for a particle initially at site $|n\rangle, n\in\mathbf{Z}^D$, to stay at this same site $|n\rangle$ the time evolution being governed only by the restriction of the Hamiltonian H_ω to the interval Δ. More precisely

$$A_n(\Delta,\omega)=\lim_{T\to\infty}\int_0^T\frac{dt}{T}\,|\langle n|\pi_\omega(e^{\imath Ht}P_\Delta)|n\rangle|^2.$$

The covariance of the Hamiltonian leads to $A_n(\Delta,\omega)=A_0(\Delta,T^{-n}\omega)$. After averaging over the disorder configurations, it will therefore be sufficient to consider only $A(\Delta,\omega)=A_0(\Delta,\omega)$. We introduce the (disorder average) return probability:

$$\xi(\Delta)=\int d\mathbf{P}(\omega)A(\Delta,\omega).$$

$\xi(\Delta)$ can also be expressed as follows

$$\xi(\Delta)=\lim_{T\to\infty}\int_0^T\frac{dt}{T}\int\frac{d^D\theta}{(2\pi)^D}\,\mathcal{T}((e^{-\imath Ht}P_\Delta)e^{\imath\theta\cdot\nabla}(e^{\imath Ht}P_\Delta)),$$

as can be checked by a direct calculation. Here $\nabla=(\partial_1,...,\partial_D)$ is the derivation on $\mathscr{A}$. If $\xi(\Delta)>0$, we expect that there are some localized states corresponding to energies within the interval Δ, this is contained in Theorem 12 below. Using Wiener's criterion we obtain

$$\xi(\Delta)=\int d\mathbf{P}(\omega)\lim_{T\to\infty}\int_0^T\frac{dt}{T}\left|\int\langle 0|\pi_\omega(dP_E P_\Delta)|0\rangle e^{\imath Et}\right|^2$$

$$=\int d\mathbf{P}(\omega)\sum_{E\in\sigma_{pp}(\omega)\cap\Delta}|\psi_{\omega,E}(0)|^4, \tag{59}$$

here $P_E\in\mathscr{W}$ is the spectral projection of H on the interval $(-\infty,E]$, $\psi_{\omega,E}$ is the eigenstate of H_ω corresponding to the eigenvalue E and $\sigma_{pp}(\omega)$ is the set of eigenvalues of H_ω. Hence we see that $\xi(\Delta)$ is related to the so-called "inverse participation ratio." The main result is summarized in the following:

Theorem 12: *Let $\Delta\subset\mathbf{R}$ be an open interval. Then the following results hold:*

(i) *Let $\mathscr{N}$ be the DOS of H. If λ is a growth point of $\mathscr{N}$ then $\lambda\in\sigma(\omega)$ **P**-almost surely. If λ is not a growth point of $\mathscr{N}$ then λ is outside $\sigma(\omega)$ **P**-almost surely.*

(ii) *The number of eigenvalues of H_ω counted with their multiplicity and contained in Δ is either zero or infinity **P**-almost surely.*

(iii) *A given real number λ is **P**-almost never an eigenvalue of H_ω with finite multiplicity.*

(iv) *The number of eigenvalues in Δ is infinite if and only if $\xi(\Delta)>0$.*

Proof: (i) Suppose λ is a growth point of $\mathscr{N}$. This implies that $\langle 0|\pi_\omega(P_\Delta)|0\rangle>0$ on a set of positive measure for any neighborhood Δ of λ. Thus, for any such neighborhood, the set of ω's such that $\sigma(\omega)\cap\Delta\neq\varnothing$ has probability one, because it is measurable and translation invariant. Since

$$\{\omega\in\Omega|\lambda\in\sigma(\omega)\}=\bigcap_{j>0}\left\{\omega\in\Omega|\sigma(\omega)\cap\left(\lambda-\frac{1}{j},\lambda+\frac{1}{j}\right)\neq\varnothing\right\},$$

then $\lambda\in\sigma(\omega)$ **P**-almost surely.

Conversely let us suppose that λ is not a growth point of $\mathscr{N}$. Then there is a small open interval Δ containing it such that $\langle 0|\pi_\omega(P_\Delta)|0\rangle=0$ on a set Ω_0 of probability one. We set $\Omega_\infty=\cap_{n\in\mathbf{Z}^D}T^{-n}\Omega_0$ to get a translation invariant subset of probability one. On this subset $\langle n|\pi_\omega(P_\Delta)|n\rangle=0$ for all $n\in\mathbf{Z}^D$ by covariance. Therefore on this subset, $\sigma(\omega)\cap\Delta=\varnothing$. This implies that λ is outside $\sigma(\omega)$ with probability one.

(ii) Let $P_{pp}(\omega)$ be the projection onto the subspace spanned by the eigenvectors of H_ω. It defines a covariant, measurable family of bounded operators.[31] So by Theorem 6 it defines an element P_{pp} of $\mathscr{W}$. Thus the number of eigenvalues counted with their multiplicity and contained in the Borel set Δ is the dimension of $\pi_\omega(P_{pp}P_\Delta)$, namely, it is zero or infinity **P**-almost surely by Proposition 11.

(iii) $\lambda$is an eigenvalue of H_ω of finite multiplicity if and only if $0<\mathrm{Tr}(\pi_\omega(P_\lambda))<\infty$. By Proposition 11 this happens with probability zero.

(iv) Let Ω_Δ be the set $\{\omega\in\Omega|\sigma_{pp}(\omega)\cap\Delta\neq\varnothing\}$. Thanks to Eq. (59) this set has positive probability if and only if $\xi(\Delta)>0$. Since this set is translation invariant and measurable it has probability one if and only if $\xi(\Delta)>0$. By Proposition 11 again, the result is achieved. □

Remark 1: (ii) says that, with probability one, no eigenvalue of H_ω with finite multiplicity is isolated, whereas (iii) shows that such eigenvalues are fluctuating with the disorder. There are however examples of models having nonfluctuating eigenvalues of infinite multiplicity.

Remark 2: The criterion (iv) does not eliminate the occurrence of some continuous spectrum.

B. Noncommutative localization criterion and localization length

We introduce a second physical idea of localization: the average mean square displacement $\delta X(T)$. Let X be the position operator in G and $X_\omega(t)=e^{iH_\omega t}Xe^{-iH_\omega t}$ its time evolution. Then we consider:

$$\delta X_{\omega,n}(T)^2=\int_0^T \frac{dt}{T}\,\langle n|(X_\omega(t)-X)^2|n\rangle.$$

The covariance relation leads to $\delta X_{\omega,n}(T)=\delta X_{T-n_\omega,0}(T)$, so that it is sufficient to examine the behavior of the state $|0\rangle$. Averaging over the disorder and using again our algebra we get

$$\delta X(T)^2=\int_0^T \frac{dt}{T}\,\mathcal{T}(|\nabla(e^{-(\iota/\hbar)Ht})|^2). \tag{60}$$

Boundedness of $\delta X(T)$ in time T will be an indicator for localization and behavior proportional to $T^\sigma, \sigma\in(0,1]$ will be interpreted as diffusive or ballistic quantum motion. In order to localize in energy we restrict the motion to energies in a Borel subset Δ of the real line and get $\delta X_\Delta(T)$ in much the same way. On the other hand taking $T\to\infty$ leads us to define the Δ-localization length as

$$l^2(\Delta)=\limsup_{T\to\infty}\int_0^T \frac{dt}{T}\,\mathcal{T}(|\nabla(e^{-(\iota/\hbar)Ht}P_\Delta)|^2). \tag{61}$$

Boundedness of $l^2(\Delta)$ will be our localization condition for the spectral subset Δ.

Theorem 13: *Suppose that $l^2(\Delta)<\infty$. Then H_ω has pure-point spectrum in Δ for almost every $\omega\in\Omega$. Moreover, if $\mathcal{N}(E)=\mathcal{T}(P_E)$ is the density of states, there is a $\mathcal{N}$-measurable function l on Δ such that for every Borel subset Δ' of Δ:*

$$l^2(\Delta')=\int_{\Delta'} d\mathcal{N}(E)l(E)^2. \tag{62}$$

Finally, if $l^2(\Delta)<\infty$, $l^2(\Delta')$ is given by

$$l^2(\Delta')=\int d\mathbf{P}(\omega)\sum_{n\in\mathbf{Z}^2}|n|^2\sum_{E\in\sigma_{pp}(\omega)\cap\Delta'}|\langle 0|\pi_\omega(P_{\{E\}})|n\rangle|^2. \tag{63}$$

Theorem 14: *If the localization condition $l^2(\Delta)<\infty$ is satisfied, then at every regularity point of the density of states, the application $E\in\Delta\to P_E\in\mathcal{S}$ is continuous.*

Remark 1: Letting Δ shrink to one point E, the function $l^2(E)$ represents a kind of average of the quantity $\Sigma_{n\in\mathbf{Z}^D}|n|^2|\psi_{E,\omega}(n)|^2$, where $\psi_{E,\omega}$ is an eigenvector of H_ω corresponding to energy E; here the average is taken over the disorder and a small spectral set around E. This quantity measures the extension of this eigenstate. For this reason the function l will be called localization length for H. Note that no exponential decay of the wave functions is needed for our localization length to be finite. However, such behavior may be studied within the present framework; we postpone the details to future work.

Remark 2: The index theorem which we proved in Sec. IV G only requires that P_E be in $\mathscr{I}$ in order to insure the integrality of the corresponding Chern character (compare to results of Sec. IV D). Thus boundedness of $l^2(\Delta)$ is sufficient but not necessary to prove the index theorem.

Remark 3: The localization condition $l^2(\Delta)<\infty$ implies pure-point spectrum in Δ almost surely. In mathematical physics, pure-point spectrum of a Hamiltonian H in a certain region of the density of states has been considered as the criterion for localization. The eigenstates being square integrable are thus localized. The classical RAGE-theorem permits to make such a statement more accurate.

Proof of Theorem 13: (i) The basic argument we will use here is due to Guarneri.[83] The spectral projection on the continuous part of the spectrum is then given by $\pi_\omega(P_c)=1-\pi_\omega(P_{pp})$. Using the definition of the trace, then applying the theorems of Fubini and monotone convergence, we find (with the notation $|n|_\infty=\max_{1\leq j\leq D}|n_j|$):

$$\delta X_\Delta(T)^2=\int d\mathbf{P}(\omega)\lim_{N\to\infty}\sum_{|n|_\infty<N}|n|^2 p_T(\omega,n),$$

where

$$p_T(\omega,n)=\int_0^T\frac{dt}{T}\,|\langle 0|\pi_\omega(e^{iHt}P_\Delta)|n\rangle|^2.$$

Then $p_T(\omega,n)$ satisfies

$$0\leq p_T(\omega,n)\leq 1,\quad \sum_{n\in\mathbf{Z}^D}p_T(\omega,n)=\langle 0|\pi_\omega(P_\Delta)|0\rangle\leq 1.$$

Now, we use the spectral theorem for H_ω and the Wiener criterion

$$\lim_{T\to\infty}p_T(\omega,n)=\lim_{T\to\infty}\int_0^T\frac{dt}{T}\left|\int\langle 0|\pi_\omega(dP_E P_\Delta)|n\rangle e^{iEt}\right|^2$$

$$=\sum_{E\in\sigma_{pp}(\omega)\cap\Delta}|\langle 0|\pi_\omega(P_{\{E\}})|n\rangle|^2.$$

In particular, for fixed positive integer L:

$$\lim_{T\to\infty}\sum_{|n|<L}p_T(\omega,n)=\sum_{E\in\sigma_{pp}(\omega)\cap\Delta}\sum_{|n|<L}\langle 0|\pi_\omega(P_{\{E\}})|n\rangle\langle n|\pi_\omega(P_{\{E\}})|0\rangle$$

$$\leq\langle 0|\pi_\omega(P_\Delta)|0\rangle-\langle 0|\pi_\omega(P_c P_\Delta)|0\rangle. \tag{64}$$

Let us now introduce the following notations:

$$p_T(n)=\int d\mathbf{P}(\omega)p_T(\omega,n),$$

$$r=\int d\mathbf{P}(\omega)\langle 0|\pi_\omega(P_c P_\Delta)|0\rangle\quad s=\int d\mathbf{P}(\omega)\langle 0|\pi_\omega(P_\Delta)|0\rangle.$$

By the dominated convergence theorem it gives

$$\int d\mathbf{P}(\omega) \lim_{T\to\infty} \sum_{|n|<L} p_T(\omega,n) = \lim_{T\to\infty} \sum_{|n|<L} p_T(n) \leqslant s-r.$$

Since $r \geqslant 0$, one can find $T_L > 0$ such that, if $T \geqslant T_L$, one has $\Sigma_{|n|<L}\, p_T(n) \leqslant s - r/2$. Thus, for $T \geqslant T_L$:

$$\delta X_\Delta(T)^2 \geqslant \int d\mathbf{P}(\omega) \sum_{|n|\geqslant L} |n|^2 p_T(\omega,n) \geqslant \frac{L^2 r}{2}. \tag{65}$$

Taking the lim sup over T we get $Lr \leqslant l^2(\Delta)$ for all L, implying that $r=0$. Thus for almost all $\omega \in \Omega$: $\langle 0|\pi_\omega(P_c P_\Delta)|0\rangle = 0$. Using the covariance relation and since $\mathbf{Z}^D$ is countable, there is $\Omega' \subset \Omega$ of probability one such that $\langle n|\pi_\omega(P_c P_\Delta)|n\rangle = 0$ for all $n \in \mathbf{Z}^D$. In particular $\pi_\omega(P_c)|n\rangle = 0$ for all $n \in \mathbf{Z}^D$ and all $\omega \in \Omega'$, namely the continuous spectrum in Δ is empty.

(ii) Given two Borel subsets $\Delta_1, \Delta_2 \subset \Delta$, we define the following expression:

$$\mathscr{E}^{(L)}_{T,\omega}(\Delta_1,\Delta_2) = \int_0^T \frac{dt}{T} \sum_{|n|<L} |n|^2 \langle 0|\pi_\omega(e^{iHt} P_{\Delta_1})|n\rangle \overline{\langle 0|\pi_\omega(e^{iHt} P_{\Delta_2})|n\rangle}.$$

This expression gives a Borel function in ω. We use the Wiener criterion:

$$\lim_{T\to\infty} \mathscr{E}^{(L)}_{T,\omega}(\Delta_1,\Delta_2) = \sum_{|n|<L} |n|^2 \sum_{E\in\sigma_{pp}(\omega)\cap\Delta_1\cap\Delta_2} |\langle 0|\pi_\omega(P_{\{E\}})|n\rangle|^2 = \mathscr{E}^{(L)}_\omega(\Delta_1\cap\Delta_2).$$

From this definition of $\mathscr{E}^{(L)}_\omega(\Delta')$ for a Borel set $\Delta' \subset \Delta$ it follows that

(a) $0 \leqslant \mathscr{E}^{(L)}_\omega(\Delta') \leqslant L^2 \langle 0|\pi_\omega(P_{pp} P_{\Delta'})|n\rangle \leqslant L^2$;
(b) if $\Delta_1 \cap \Delta_2 = \varnothing$, then $\mathscr{E}^{(L)}_\omega(\Delta_1 \cup \Delta_2) = \mathscr{E}^{(L)}_\omega(\Delta_1) + \mathscr{E}^{(L)}_\omega(\Delta_2)$;
(c) $\mathscr{E}^{(L)}_\omega(\Delta') \leqslant \mathscr{E}^{(L+1)}_\omega(\Delta')$;
(d) $\mathscr{E}^{(L)}_\omega(\Delta')$ is a Borel function of ω as pointwise limit of Borel functions;
(e) if $(\Delta_j)_{j\in\mathbf{N}}$ is a decreasing sequence of Borel subsets such that $\cap_{j\in\mathbf{N}} \Delta_j = \varnothing$, then $\mathscr{E}^{(L)}_\omega(\Delta_j)$ decreases to zero.

Averaging over the disorder, we obtain $\mathscr{E}^{(L)}(\Delta') = \int d\mathbf{P}(\omega)\, \mathscr{E}^{(L)}_\omega(\Delta')$ which fulfills (a),(b),(c), and also (e) thanks to the dominated convergence theorem. Moreover, since $\mathscr{E}^{(L)}(\Delta') \leqslant L^2$, we can use the dominated convergence theorem, Fubini's theorem, and the definition of $l^2(\Delta)$ to get

$$\mathscr{E}^{(L)}(\Delta') = \lim_{T\to\infty} \int_0^T \frac{dt}{T} \int d\mathbf{P}(\omega) \sum_{|n|<L} |n|^2 |\langle 0|\pi_\omega(e^{iHt} P_{\Delta'})|n\rangle|^2 \leqslant l^2(\Delta').$$

In much the same way, we get, thanks to (b), $\mathscr{E}^{(L)}(\Delta') \leqslant l^2(\Delta) < \infty, \forall L \in \mathbf{N}$. As $\mathscr{E}^{(L)}(\Delta')$ is bounded and increasing in L, it follows that $\mathscr{E}(\Delta') = \lim_{L\to\infty} \mathscr{E}^{(L)}(\Delta')$ exists. $\mathscr{E}$ defines a nonnegative set function on the set of Borel subsets of Δ. Because of property (e), it is continuous from above and since it is moreover finite, we can conclude its σ-additivity. Therefore $\mathscr{E}$ is a Radon measure. Moreover, using the monotone convergence theorem:

$$\mathscr{E}(\Delta') = \int d\mathbf{P}(\omega) \sum_{E\in\sigma_{pp}(\omega)\cap\Delta'} \sum_{n\in\mathbf{Z}^D} |n|^2 |\langle 0|\pi_\omega(P_{\{E\}})|n\rangle|^2.$$

(iii) Let us show now $l^2(\Delta') \leqslant \mathscr{E}(\Delta')$. Recall that

$$\mathscr{E}(\Delta')=\int d\mathbf{P}(\omega)\sup_{N\in\mathbf{N}}\sum_{|n|<N}|n|^2\lim_{T\to\infty}\int_0^T\frac{dt}{T}|\langle 0|\pi_\omega(e^{iHt}P_{\Delta'})|n\rangle|^2<\infty.$$

We can replace the limit by the lim sup. Hence for a fixed T_0 large enough, the following expression is finite and we may apply Fubini's theorem:

$$\begin{aligned}&\int d\mathbf{P}(\omega)\sup_{N\in\mathbf{N}}\sum_{|n|<N}|n|^2\sup_{T>T_0}\int_0^T\frac{dt}{T}|\langle 0|\pi_\omega(e^{iHt}P_{\Delta'})|n\rangle|^2\\ &\geqslant\sup_{T>T_0}\int d\mathbf{P}(\omega)\int_0^T\frac{dt}{T}\sup_{N\in\mathbf{N}}\sum_{|n|<N}|n|^2|\langle 0|\pi_\omega(e^{iHt}P_{\Delta'})|n\rangle|^2\\ &\geqslant\limsup_{T>T_0}\int_0^T\frac{dt}{T}\int d\mathbf{P}(\omega)\sup_{N\in\mathbf{N}}\sum_{|n|<N}|n|^2|\langle 0|\pi_\omega(e^{iHt}P_{\Delta'})|n\rangle|^2\\ &=l^2(\Delta').\end{aligned}$$

As this is true for all T_0, we obtain $l^2(\Delta')\leqslant\mathscr{E}(\Delta')$, and with (ii) their equality.

(iv) To finish the proof we use the Radon–Nikodym theorem. It is thus sufficient to show that the measure $l^2(\Delta')=\mathscr{E}(\Delta')$ is absolutely continuous with respect to the density of states. Let $\Delta'\subset\Delta$ be such that $\mathscr{N}(\Delta')=\mathscr{T}(P_{\Delta'})=0$. From the definition of the trace it follows that $\langle 0|\pi_\omega(P_{\Delta'})|0\rangle=0$ almost surely. By covariance and because $\mathbf{Z}^D$ is countable, this gives $\pi_\omega(P_{\Delta'})|n\rangle=0$ for all $n\in\mathbf{Z}^D$ almost surely. Then the definition of $\mathscr{E}^{(L)}_{T,\omega}(\Delta',\Delta')$ implies that it is zero for any l,T and almost all ω; consequently $0=\mathscr{E}(\Delta')=l^2(\Delta')$. □

Let us give another useful expression for $l^2(\Delta)$. We consider finite partitions $\mathscr{P}$ of Δ into disjoint Borelian subsets:

$$\mathscr{P}=\{\Delta_j\subset\Delta,\Delta_j\ \text{Borelian}\mid j=1\ldots q,\cup_j\Delta_j=\Delta,\Delta_j\cap\Delta_k=\varnothing\}.$$

The set $\mathscr{Z}(\Delta)$ of such finite partitions is ordered by refinement

$$\mathscr{P}\leqslant\mathscr{P}'\quad\Leftrightarrow\quad\forall\Delta'\in\mathscr{P}'\,\exists\Delta''\in\mathscr{P}\ \text{ such that }\ \Delta'\subset\Delta''.$$

This gives an ordered net. We define

$$\hat{l}^2(\Delta)=\lim_{\mathscr{P}\in\mathscr{Z}(\Delta)}\sum_{\Delta'\in\mathscr{P}}\mathscr{T}(|\nabla P_{\Delta'}|^2),\tag{66}$$

where the limit is understood to be the one under refinements of partitions.

Proposition 12: $\hat{l}^2(\Delta)$ *is well defined in* $\bar{\mathbf{R}}=\mathbf{R}\cup\{\infty\}$. *Moreover:*

$$\hat{l}^2(\Delta)=\sup_{\mathscr{P}\in\mathscr{Z}(\Delta)}\sum_{\Delta'\in\mathscr{P}}\mathscr{T}(|\nabla P_{\Delta'}|^2).$$

Proof: It is sufficient to show that refining the partition results in the increase of the quantity $\Sigma_{\Delta'\in\mathscr{P}}\mathscr{T}(|\nabla P_{\Delta'}|^2)$. Take $\Delta=\Delta_1\cup\Delta_2,\Delta_1\cap\Delta_2=\varnothing$, and suppose $\mathscr{T}(|\nabla P_\Delta|^2)<\infty$. We have

$$\mathscr{T}(|\nabla P_\Delta|^2)=\mathscr{T}(|\nabla P_{\Delta_1}|^2+|\nabla P_{\Delta_2}|^2+\nabla P_{\Delta_1}\cdot\nabla P_{\Delta_2}+\nabla P_{\Delta_2}\cdot\nabla P_{\Delta_1}).$$

Now, either one of the Sobolev norms of P_{Δ_1} and P_{Δ_2} is infinite and in this case the inequality $\mathcal{T}(|\nabla P_{\Delta_1\cup\Delta_2}|^2)\ \mathcal{T}(|\nabla P_{\Delta_1\cup\Delta_2}|^2)\leq\mathcal{T}(|\nabla P_{\Delta_1}|^2)+\mathcal{T}(|\nabla P_{\Delta_2}|^2)$ is trivially satisfied, or the components of ∇P_{Δ_1} and ∇P_{Δ_2} are in $L^2(\mathcal{A},\mathcal{T})$. The Hölder inequality implies $\nabla P_{\Delta_1}\cdot\nabla P_{\Delta_2}\in L^1(\mathcal{A},\mathcal{T})$. We may therefore treat each term separately.

$$\mathcal{T}(\nabla P_{\Delta_1}\cdot\nabla P_{\Delta_2})=\mathcal{T}(P_{\Delta_1}\nabla P_{\Delta_1}(1-P_{\Delta_1})\cdot\nabla P_{\Delta_2})+\mathcal{T}((1-P_{\Delta_1})\nabla P_{\Delta_1}P_{\Delta_1}\cdot\nabla P_{\Delta_2}).$$

We apply the cyclicity of $\mathcal{T}$ and the formulas

$$P_{\Delta_1}\leq 1-P_{\Delta_2},\quad P_{\Delta_2}\leq 1-P_{\Delta_1},\quad \nabla P_{\Delta_2}P_{\Delta_1}=-P_{\Delta_2}\nabla P_{\Delta_1},$$

to get

$$\begin{aligned}\mathcal{T}(\nabla P_{\Delta_1}\cdot\nabla P_{\Delta_2})&=\mathcal{T}(P_{\Delta_1}\nabla P_{\Delta_1}\cdot\nabla P_{\Delta_2})+\mathcal{T}(\nabla P_{\Delta_1}\cdot(-\nabla P_{\Delta_1}P_{\Delta_2}))\\&=-\mathcal{T}(|\nabla P_{\Delta_2}P_{\Delta_1}|^2)-\mathcal{T}(|\nabla P_{\Delta_1}P_{\Delta_2}|^2)\leq 0.\end{aligned}$$

The same calculation for the other term then implies the result. □

Because of Eq. (63), one expects $\hat{l}^2$ to be equal to l^2 and this is what we shall prove in the sequel. We first need the following technical lemma:

Lemma 16: Let μ,ν be two finite positive measures on a Borel subset Δ of the real line and μ_{pp},ν_{pp} their pure-point parts. For a finite partition $\mathcal{P}$ of Δ we set

$$\mathcal{N}(\Delta,\mathcal{P})=\sum_{\Delta'\in\mathcal{P}}\mu(\Delta')\nu(\Delta').$$

Then

$$\lim_{\mathcal{P}\in\mathcal{Z}(\Delta)}\mathcal{N}(\Delta,\mathcal{P})=\inf_{\mathcal{P}\in\mathcal{Z}(\Delta)}\mathcal{N}(\Delta,\mathcal{P})=\sum_{E\in\Delta}\mu_{pp}(\{E\})\nu_{pp}(\{E\}).$$

Proof: First, we show that $\mathcal{N}(\Delta,\mathcal{P})$ decreases as the partition is refined. We introduce $\mathcal{F}(\mathcal{P})=\cup_{\Delta'\in\mathcal{P}}\Delta'\times\Delta'$. Suppose that $\mathcal{P}\leq\mathcal{P}'$, then $\mathcal{F}(\mathcal{P}')\subset\mathcal{F}(\mathcal{P})$. As $\mu\otimes\nu$ is a finite positive measure on a Borel subset of $\mathbf{R}^2$, we have

$$\mathcal{N}(\Delta,\mathcal{P})=\mu\otimes\nu(\mathcal{F}(\mathcal{P}))\geq\mu\otimes\nu(\mathcal{F}(\mathcal{P}'))=\mathcal{N}(\Delta,\mathcal{P}').$$

Let us now consider the Lebesgue decomposition of the measures μ and ν in continuous and pure-point parts:

$$\mu\otimes\nu=\mu_{pp}\otimes\nu_{pp}+\mu_c\otimes\nu_{pp}+\mu_{pp}\otimes\nu_c+\mu_c\otimes\nu_c.$$

As the measures are finite, the continuous parts satisfy

$$\forall\epsilon>0\,\exists\,\delta>0\quad\text{such that}\quad\mu_c(\Delta')<\epsilon,\quad\nu_c(\Delta')<\epsilon\,\forall\Delta'\quad\text{with diam}(\Delta')<\delta.$$

Here, the diameter is defined by $\text{diam}(\Delta')=\sup_{x,y\in\Delta'}|x-y|$. Now we choose and fix a sequence of $\mathcal{P}_n$ of finite partitions which satisfies:

$$\lim_{n\to\infty}(\max_{\Delta'\in\mathcal{P}_n}\text{diam}(\Delta'))=0.$$

For such a sequence $\mathscr{P}_n$ the following holds:

$$\bigcap_{n=1}^{\infty}\Big(\bigcup_{\Delta'\in\mathscr{P}_n}(\Delta'\times\Delta')\Big)=\{(x,x)|x\in\Delta\}=\mathrm{Diag}(\Delta\times\Delta).$$

In the limit, the contribution containing a continuous part vanishes for indeed for instance

$$\lim_{n\to\infty}\mu_c\otimes\nu_{pp}\Big(\bigcup_{\Delta'\in\mathscr{P}_n}(\Delta'\times\Delta')\Big)\leqslant\lim_{n\to\infty}\epsilon(n)\sum_{\Delta'\in\mathscr{P}_n}\nu_{pp}(\Delta')\leqslant\lim_{n\to\infty}\epsilon(n)\nu(\Delta)=0.$$

Therefore, using the σ-additivity of the finite measure $\mu\otimes\nu$ (which is equivalent to continuity from above), we find

$$\lim_{n\to\infty}\sum_{\Delta'\in\mathscr{P}_n}\mu(\Delta')\nu(\Delta')=\lim_{n\to\infty}\sum_{\Delta'\in\mathscr{P}_n}\mu_{pp}(\Delta')\nu_{pp}(\Delta')=\mu_{pp}\otimes\nu_{pp}\Big(\bigcap_{n=1}^{\infty}\bigcup_{\Delta'\in\mathscr{P}_n}(\Delta'\times\Delta')\Big)$$

$$=\mu_{pp}\otimes\nu_{pp}\mathrm{Diag}(\Delta\times\Delta)=\sum_{E\in\Delta}\mu_{pp}(\{E\})\otimes\nu_{pp}(\{E\}).$$

This is true for every such refining sequences of partitions, leading to the result. □

Proposition 13: If $l^2(\Delta)<\infty$, *then* $l^2(\Delta')=\hat{l}^2(\Delta')$ *for every Borelian subset* $\Delta'\subset\Delta$.

Proof: As $l^2(\Delta)<\infty$, Theorem 13 shows that

$$\mathscr{E}^{(L)}(\Delta)=\int d\mathbf{P}(\omega)\sum_{|n|<L}|n|^2\sum_{E\in\sigma_{pp}(\omega)\cap\Delta}|\langle 0|\pi_\omega(P_{\{E\}})|n\rangle|^2.$$

We decompose the complex measure $\mu(\Delta')=\langle 0|\pi_\omega(P_{\Delta'})|n\rangle$ by polarization into four positive measures:

$$\mu(\Delta')=\mu_1(\Delta')-\mu_2(\Delta')+\imath\mu_3(\Delta')-\imath\mu_4(\Delta').$$

With these notations, we apply Lemma 16 to each term to get

$$\mathscr{E}^{(L)}(\Delta)=\int d\mathbf{P}(\omega)\sum_{|n|<L}|n|^2\sum_{E\in\sigma_{pp}(\omega)\cap\Delta}\left(\sum_{k=1}^{4}\mu_k\otimes\mu_k-2\mu_1\otimes\mu_2-2\mu_3\otimes\mu_4\right)(\{E\times E\})$$

$$=\int d\mathbf{P}(\omega)\sum_{|n|<L}|n|^2\lim_{\mathscr{P}\in\mathscr{Z}(\Delta')}\sum_{\Delta''\in\mathscr{P}}|\langle 0|\pi_\omega(P_{\Delta''})|n\rangle|^2.$$

The dominated convergence theorem now gives

$$\mathscr{E}^{(L)}(\Delta)=\lim_{\mathscr{P}\in\mathscr{Z}(\Delta')}\sum_{\Delta''\in\mathscr{P}}\int d\mathbf{P}(\omega)\sum_{|n|<L}|n|^2|\langle 0|\pi_\omega(P_{\Delta''})|n\rangle|^2=\hat{\mathscr{E}}^{(L)}(\Delta),$$

by definition of $\hat{\mathscr{E}}^{(L)}(\Delta)$. Obviously we have

$$\hat{\mathscr{E}}^{(L)}(\Delta)\leqslant\lim_{\mathscr{P}\in\mathscr{Z}(\Delta')}\sum_{\Delta''\in\mathscr{P}}\int d\mathbf{P}(\omega)\sum_{n\in\mathbf{Z}}|n|^2|\langle 0|\pi_\omega(P_{\Delta''})|n\rangle|^2=\hat{l}^2(\Delta').$$

Moreover, $\hat{\mathcal{E}}^{(L)}(\Delta')$ is increasing in L. Since $\mathcal{E}^{(L)}(\Delta)$ is bounded by $l^2(\Delta)$, so is $\hat{\mathcal{E}}^{(L)}(\Delta)$. Its limit therefore exists. Actually, it converges to $\hat{l}^2(\Delta')$. For indeed, with the dominated convergence theorem, we find

$$\begin{aligned}\hat{l}^2(\Delta') &= \sup_{\mathcal{P}\in\mathcal{Z}(\Delta')} \sum_{\Delta''\in\mathcal{P}} \int d\mathbf{P}(\omega) \lim_{L\to\infty} \sum_{|n|<L} |n|^2 |\langle 0|\pi_\omega(P_{\Delta''})|n\rangle|^2 \\ &= \sup_{\mathcal{P}\in\mathcal{Z}(\Delta')} \sup_{L\in\mathbf{N}} \sum_{\Delta''\in\mathcal{P}} \int d\mathbf{P}(\omega) \sum_{|n|<L} |n|^2 |\langle 0|\pi_\omega(P_{\Delta''})|n\rangle|^2 \\ &= \lim_{L\to\infty} \hat{\mathcal{E}}^{(L)}(\Delta').\end{aligned}$$

This finishes the proof. □

Proof of Theorem 14. Fix $E,E' \in \Delta, E\leqslant E'$, then

$$\begin{aligned}\|P_{E'}-P_E\|_{\mathcal{J}}^2 = \|P_{[E,E']}\|_{\mathcal{J}}^2 &\leqslant \hat{l}^2([E,E'])+\mathcal{N}(E')-\mathcal{N}(E) \\ &= \int_E^{E'} (l(E'')^2+1)\,d\mathcal{N}(E'').\end{aligned}$$

Now, in the limit $E'\to E$, this is zero if E is a regularity point of $\mathcal{N}$. □

C. Localization in physical models

In this section we give an example of a physical model for which the localization condition $l^2(\Delta)<\infty$ is satisfied both for weak disorder at the band edges and for high disorder all over the spectrum. For the mathematical treatment we will, once again, restrict ourselves to the discrete case. Our line of arguments will use results of Aizenman and Molchanov.[84,40] They give a simple proof of mathematical results proved earlier.[35,85,38,31,37] We will conclude with some remarks about the continuous case.

As a preliminary, let us remark that if the spectrum has a finite gap Δ, then the condition $l^2(\Delta)<\infty$ is satisfied for the simple reason that $P_\Delta=0$. In dimension two, the Chern character corresponding to every energy band is therefore an integer and the Hall conductivity an integer multiple of e^2/h. However, this integer may be zero.

The analysis by Aizenman and Molchanov consists of two steps: first, one shows exponential decay of low moments of Green's function for concrete classes of models and in specified regions Δ of the spectrum; then, this decay is used[84] to show exponential decay in $|n-m|$ of the quantity $\int d\mathbf{P}(\omega)\sup_{t\geqslant 0}\langle n|\pi(e^{iHt}P_\Delta)|m\rangle$. This, in turn, will allow us to show that $l^2(\Delta)<\infty$ is satisfied.

Here, we fix our attention to the D-dimensional Anderson model with constant magnetic field. The Hamiltonian acting on $\ell^2(\mathbf{Z}^D)$ is given by

$$H_{\omega,\lambda}=H_0+\lambda V_\omega, \quad \lambda\in\mathbf{R},$$

where H_0 is the D-dimensional analog of Hamiltonian given in Eq. (10). V_ω is the disorder potential: at every site of the lattice it takes a random value in the interval $[-1,1]$; the probability distribution is supposed to be uniform and the sites are independent. In Refs. 40 and 84 much more general situations are considered: the kinetic part H_0 may have nonzero elements farther off the diagonal as long as they decay exponentially in distance from the diagonal; the random potential may also have gaussian or Cauchy distribution; moreover, correlations between the sites are allowed in a sense to be made precise. For us, however, the Anderson model captures the essential of the localization phenomenon and it has the advantage that the hull $(\Omega,T,\mathbf{P})$ is easily

constructed as topological product of intervals $[-1, 1]$; by Tychonow's theorem, Ω is compact; the action T on Ω is given by the translations in physical space and the product measure of the uniform probability on $[-1, 1]$ is a T-invariant and ergodic measure $\mathbf{P}$ on Ω.

We will now describe precisely in which situations Aizenman and Molchanov prove exponential decay of low moments of Green's functions for the Anderson model:

$$\int d\mathbf{P}(\omega)\left|\langle n|\pi_\omega\left(\frac{1}{H-E}\right)|m\rangle\right|^s \leqslant De^{-c|n-m|}, \quad s\in(0,1). \tag{67}$$

The first situation is that of strong disorder: for every $s\in(0,1)$ there exists an $\lambda_c=\lambda_c(D,s)$ such that for $\lambda\geqslant\lambda_c$ the bound holds on all of the spectral axis.

The second situation considers weak disorder, that is small λ. Recall that the free Hamiltonian H_0 has energy bands of continuous spectrum. In the interesting case for us of dimension two, the energy bands and their dependence on the magnetic field are given by Hofstadter's butterfly; H_0 then gives Harper's equation. For the Anderson model the bound (67) is obtained for energies E situated at band edges:

$$E\in R(\lambda)=\bigcup_{s\in(0,1)}\{E\in\mathbf{R}-\sigma(H_0)|\tfrac{1}{2}\lambda g_s(E)<1\},$$

where

$$g_s(E)=\sup_{x\in\mathbf{Z}^D}\left(\sum_{y\in\mathbf{Z}^D}\left|\langle n|\pi_\omega\left(\frac{1}{H-E}\right)|m\rangle\right|^s\right)^{-s}.$$

[For the Anderson model, the constant $\kappa(s)$ appearing in Ref. 84 is equal to $\frac{1}{2}$ independent of s.] By a Combes–Thomas argument, one gets[84] the asymptotic behavior of $g_s(E)$ in $\xi=\text{dist}(E,\sigma(H_0))$:

$$g_s(E)=\begin{cases} s^{-\frac{D}{s}}\xi^{-(1+D/s)}, & \xi\to 0 \\ \xi^{-1}\left(1+\mathcal{O}\left(\frac{1}{\xi}\right)\right), & \xi\to\infty \end{cases}.$$

Therefore, because $\sigma(H_\lambda)\subset\{E|E\in\{\sigma(H_o)+\nu\};\nu\leqslant\lambda\}$, the intersection of $R(\lambda)$ with $\sigma(H_\lambda)$ is nonempty for λ sufficiently big. In other words, the bound (67) is satisfied for some energies belonging to the spectrum and situated at the band edges. This is probably the case as soon as $\lambda\neq 0$; dimension two is critical.[28]

Let us now come to the second step of Aizenman's analysis. If the bound (67) is satisfied for an interval Δ of the spectral axis, he proves that the unitary evolution operator filtered with the spectral projection P_Δ satisfies:

$$\int d\mathbf{P}(\omega)\sup_{t\geqslant 0}|\langle n|\pi_\omega(P_\Delta e^{iHt})|m\rangle|\leqslant Ce^{-D|n-m|}, \quad C,D\in\mathbf{R}_+. \tag{68}$$

Let us now use this result to show that the noncommutative localization condition is satisfied. Because $\|\pi_\omega(P_\Delta e^{iHt})\|_{\mathcal{L}(\ell^2\mathbf{Z}^D)}\leqslant 1$, we have

$$\mathcal{T}(|\partial(P_\Delta e^{iHt})|^2)=\sum_{n\in\mathbf{Z}^D}|n|^2\int d\mathbf{P}(\omega)|\langle 0|\pi_\omega(P_\Delta e^{iHt})|n\rangle|^2$$

$$\leqslant \sum_{n\in\mathbf{Z}^D} |n|^2 \int d\mathbf{P}(\omega)|\langle 0|\pi_\omega(P_\Delta e^{iHt})|n\rangle|$$

$$\leqslant \sum_{n\in\mathbf{Z}^D} |n|^2 Ce^{-D|n|} \leqslant C' < \infty.$$

Now, taking the time-average we see that $l^2(\Delta)<\infty$. This is, of course, only true if either $\lambda>\lambda_c$ or $\Delta\subset R(\lambda)$.

Let us now comment on the case of continuous physical space. First of all, if the Fermi energy E_F lies in a finite gap, then it can be shown that $P_F\in\mathscr{S}$.[86] For strong disorder, one expects a finite localization length, but near the set $\{\hbar\omega_c(n+\frac{1}{2})|\omega_c=e\mathscr{B}/m, n\in\mathbf{Z}\}$ the localization length probably diverges. This would be the analog of the discrete case.

VI. APPLICATIONS AND COMPLEMENTS

A. Low-lying states do not contribute to the IQHE

In this section, we address the following question: consider the Landau Hamiltonian H_L and add a periodic potential V_p of varying strength

$$H=H_L+\sigma V_p, \quad 0\leqslant V_p\leqslant 1.$$

How do the Chern numbers evolve as the coupling parameter σ is increased? In particular, what happens if $\sigma\gg\hbar\omega_c$?

Halperin *et al.*[87] made corresponding numerical calculations. They took a finite size sample and computed the Chern numbers for various values of σ by the method described in Sec. VI C. For small σ, the weak periodic potential approximation is valid and the Chern numbers are those given by the Diophantine equation (70). For intermediate values of σ, the energy bands cross each other in a complicated way and it is difficult to follow the Chern numbers. However, for $\sigma\gg\hbar\omega_c$, Halperin *et al.* observed the following striking result: all low energy bands of the spectrum carry zero Hall current. This result was put on a rigorous basis by Nakamura and one of the authors.[86]

Theorem 15: *Consider the Hamiltonian $H=H_L+\sigma V$ where V is a potential satisfying the conditions below. Then for σ sufficiently large the spectrum has gaps between his low energy bands (energy smaller than $\epsilon\hbar\omega_c$, where ϵ appears in the conditions below). If the Fermi energy lies in one of these gaps, the Hall conductivity vanishes.*

Hypothesis on the potential V:

(i) $\inf_{x\in\mathbf{R}^2} V(x)=0$, $\sup_{x\in\mathbf{R}^2} V(x)\leqslant C<\infty$.
(ii) There is a countable set of $\mathbf{R}^2$, $\{x_n, n=1,2,...\}$ such that $|x_n-x_m|\geqslant d$ if $n\neq m$ with $d>0$.
(iii) There are $\epsilon>0$ and $\mathscr{V}\in C^2(B_\epsilon)$ where $B_\epsilon=\{x\in\mathbf{R}^2||x|<\epsilon\}$ such that $d>2\epsilon$ and $V(x_n+x)=\mathscr{V}(x)$ for $x\in B_\epsilon$ and for all n,
(iv) 0 is the unique nondegenerate minimum of $\mathscr{V}\in B_\epsilon$.
(v) If $|x_n-x|\geqslant\epsilon$ for all n, then $V(x)\geqslant\delta>0$.

Note that this potential is not strictly periodic, although the bottoms of the wells need to be identical. The framework of the proof is that of noncommutative geometry completed by semiclassical analysis. Let us just describe the strategy here. The existence of the gaps follows from semiclassical analysis: the energy levels of a single, isolated quantum well are enlarged by the tunneling effect; the band width can be estimated by $e^{-a\sqrt{\sigma}}, a>0$. In order to show that the Chern character is zero, one proceeds as follows. First consider the situation of quantum wells separated

by infinitely high barriers. The projection P_E^j on the energy level of quantum well number j is one-dimensional. Its Chern character therefore vanishes. Now we consider the projection $P_\infty = \oplus_j P_E^j$, where the sum runs over all wells. Because of the additivity of the Chern character (compare Lemma 12), the Chern character of P_∞ is zero as well. Finally, semiclassical analysis allows us to show that P_∞ is Murray—von Neumann equivalent to the projection P_E onto the energy band which arises as the barriers are reduced to finite height. In view of Lemma 11, this finishes the proof. It is likely that the result also holds if the projection is not in the gaps.

B. Where and how does the localization length diverge?

In this section, we review the most important results on the behavior of the localization length at the center of the broadened Landau level. Discussion of theoretical ideas is followed by a brief presentation of numerical techniques and results. As before, we consider a system of independent fermions in a disorder potential described by a one-particle Hamiltonian.

In numerical calculations and scaling theory, a rather strong notion of localization is often used, that of a Lyapunov exponent $\lambda(E)$. If one supposes exponential decay of the correlation function of the wave function ψ_E corresponding to the energy E, $\lambda(E)$ is defined by

$$\langle \overline{\psi_E(r)}\psi_E(r')\rangle \propto e^{-\lambda(E)|r-r'|}.$$

Here the average is taken over disorder configurations. Physicists often call the inverse of $\lambda(E)$ *the localization length.* However, except for in the one-dimensional case, there is yet no clear mathematical formulation.

In 1979, the gang of four[29] used scaling theory (ideas due to Thouless [41]) and renormalization group calculations to show that, at absence of magnetic field, two is the critical dimension for localization in the following sense: for dimension smaller than two, all states are exponentially localized at any strength of disorder, that is, for all wave functions there are positive Lyapunov exponents. For dimension greater than two, there exist extended states for low disorder. At dimension two, all states are localized except for states corresponding to isolated critical energies at which the localization length diverges. In the case of perpendicular magnetic field in two dimension, the same result holds as shows a development of higher order.[42]

In the QHE, the Hall conductivity does not vanish. This led Halperin to postulate the existence of extended states.[51] A corollary of Theorem 1 is the following:

Corollary 13:[13] *Between different plateaux, there has to be a spectral interval* Δ *such that the localization condition* $l^2(\Delta)<\infty$ *is violated.*

Remark that the interval may be very small. It is even likely that there is, in fact, a single energy E such that $P_{\leqslant E}$ is not in the Sobolev space $\mathscr{S}$. This implies, of course, $l^2(\Delta)=\infty$ for every interval containing E. Corresponding numerical results are presented at the end of this section.

We now come to the question of how the localization length diverges. If one considers a disorder potential varying on a length scale much bigger than the magnetic length, the motion can semiiclassically (in the magnetic field) be well approximated by a motion along equipotentials of the disorder potential. Percolation theory now indicates that there is one critical energy E_c to which correspond extended states. Moreover, at this energy the localization length diverges as

$$\frac{1}{\lambda(E)} \propto |E-E_c|^{-\nu}. \tag{69}$$

Trugman evaluated $\nu=\frac{4}{3}$,[88] but more thorough examination of the effect of tunneling between the wells and mountains led Mil'nikov and Sokolov to predict $\nu=\frac{7}{3}$.[89] In spite of the simplicity of

the theoretical approach, the agreement with numerical results obtained by three different methods is very good. These latter are the Thouless number method, the Chern number method and the third one is a scaling theory approach; the original ideas behind are due to Refs. 90, 91, 43, and 44, respectively.

The Thouless number method is the most elementary because only energies and no eigenvectors need to be calculated. In the interesting case of strong magnetic field, calculations were made by Ando.[92] The Chern number method[93] is described in some more detail in the next section. Let us concentrate on the scaling theory approach. One considers a cylinder of (small) circumference M and extended to infinity in one direction (denoted by the variable z). The magnetic field is perpendicular to the cylinder surface. Choosing the Landau gauge $A=(0,0,\mathscr{B}z)$, the eigenstates of the free Hamiltonian are centered on circles around the cylinder corresponding to fixed values of z. This discrete set of eigenstates will serve as base for the Hilbert space; in this base one also expresses Hamiltonian perturbed by disorder. Now for fixed circumference and fixed energy, one calculates numerically the disorder averaged Lyapunov exponent L_M of a wave function by the transfer matrix method. Finite size scaling theory[43] then allows to estimate the behavior of the localization length near the critical energy; in particular, the exponent ν in Eq. (69) can be calculated. In presence of magnetic field, this method was first used in Refs. 45 and 48; more recent and more extended numerical studies[46,94] confirm the theoretical prediction $\nu=\frac{7}{3}$. It seems that the exponent is independent of the disorder distribution and the Landau level index (if the Landau level interaction are taken into account).

What is the behavior of the wave function at the critical energy E_c itself? Pichard and Sarma[43] suggest the following in order to calculate the exponent x for the powerlaw decay of the wave function ψ_{E_c}:

$$|\psi_{E_c}(r)\psi_{E_c}(r')|\propto|r-r'|^{-x}.$$

One calculates as before the Lyapunov exponent L_M for different M at fixed energy E_c. With the hypothesis of conformal invariance, x can then be shown to be equal to the slope of the linear relation between M and $1/L_M$. Chalker and Coddington[95] estimate $x\approx 0.27$. We remark that this implies that the weak localization condition $P_{\leqslant E_c}\in\mathscr{S}$ is violated. Therefore, at critical energies (that is at the centers of the broadened Landau levels) the Chern number may change.

C. Chern numbers and localization in Harper's equation

The Chern characters or Chern numbers of the energy bands of Harper's equation corresponding to rational flux have already been computed by Thouless, Kohmoto, Nightingale, and den Nijs;[8] we gave some insight in Sec. II F. Ando[96,92] and more recently Tan[97] have made numerical studies on the influence of an added disorder potential on these Chern numbers. We describe some of the results because they constitute a good illustration of theoretical ideas.

Let us first recall that Harper's equation (11) is obtained in two different limiting cases. One consists in adding a weak periodic potential V_p to the Landau Hamiltonian H_L, for example, a two-dimensional sinusoide; projecting H_L+V_p on a subspace corresponding to one Landau level then leads to Harper's equation. On the other hand, if one adds a strong periodic potential (which means that its well is much deeper than $\hbar\omega_c$), the low energy region of the spectrum is well described by a tight binding model [Eq. (10)]. As shown in Sec. II F, Bloch's representation also leads to Harper's equation.

The spectrum of Harper's equation is well known and its dependence on the magnetic flux per unit area ϕ is represented in Hofstadter's butterfly.[99] The parameter α appearing in Eq. (11) is proportional to ϕ in the tight-binding case and proportional to the inverse of ϕ in the case of weak periodic potential. Let us now suppose the rationality of $\alpha=q/p$, where p and q are relatively prime integers. The spectrum then consists of p energy bands of extended states. The Chern

number of the rth band can be calculated with the Diophantine equation in the following way.[8,100] There exist unique integers s_r, t_r (except for even p and $r=p/2$) such that $|s_r| \leq p/2$ and satisfying

$$s_r q + t_r p = r \quad r = 0, \ldots, p. \tag{70}$$

The Chern number of the rth band and therefore its contribution in units of e^2/h to the Hall conductance is given by $t_r - t_{r-1}$ in the weak periodic potential case and by $s_r - s_{r-1}$ for the tight-binding model. Remark that as $t_0 = 0$ and $t_p = 1$, the sum of all p Chern numbers in the weak periodic potential limit is equal to one. This reflects the fact that the bands result from the splitting of one Landau level with Chern number 1. On the other hand, $s_0 = s_p = 0$, such that the sum of the Chern numbers in the tight binding approximation is zero.

Now a disorder potential will be added. This could be, for example, an Anderson-type perturbation (compare Sec. V C), but in the numerical simulations a densely distributed potential with two values $V_d, -V_d$ was used. The fractal structure of Hofstadter's butterfly disappears as soon as disorder is turned on. Because many gaps close, it becomes possible to label the remaining energy bands by a finite number of rationals for any (maybe nonrational) parameter α. What are the Chern numbers of the bands and how do they evolve as a function of the disorder strength?

Before starting the discussion, let us comment on the physical relevancy of the two approximations. In the usual QHE samples presented in Sec. II B, neither seems to describe the reality: the disorder potential is much stronger than the weak periodic potential due to the ions within the surface. On the other hand, because the mobility is very high, the tight-binding approximation is not appropriate. However, lately so-called antidot superlatices seem to open the possibility of the experimental realization of parts of Hofstadter's spectrum (of course, only in a very approximative way). In these systems the ideas we present here could be tested.

For all details on how the numerical results were obtained, we refer to the original works,[96,97] but let us describe the principal steps. Of course, the calculations are made in finite size samples and therefore disorder is just treated in an increased unit cell, causing thus a splitting of the Harper

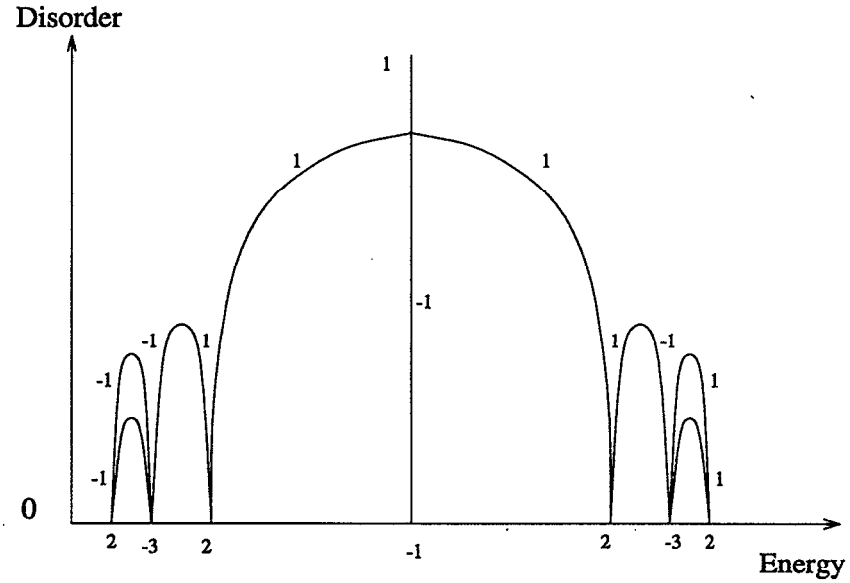

FIG. 5. Suggested phase diagram for the weak potential approximation.

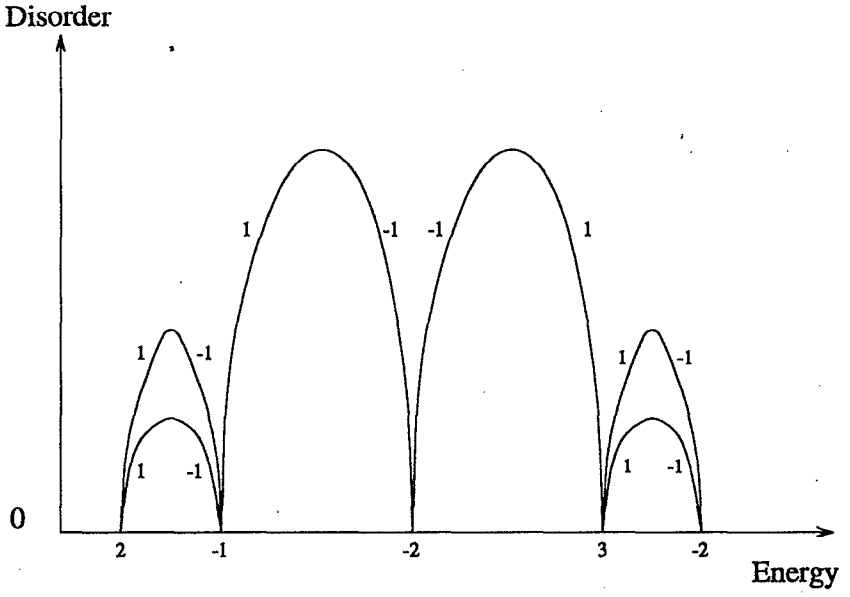

FIG. 6. Suggested phase diagram for the tight-binding approximation.

bands corresponding to a fixed rational flux into subbands. Chern numbers for these subbands can be calculated by the contour integrals given in Eq. (15). Remark that because we are in dimension two, these subbands have corresponding Hilbert subspaces generically well separated except for in isolated points. If this happens, the Chern number is simply calculated for a two dimensional fiber bundle (this is also the case for the closed central gaps of the Harper equation with even denominator). In this way, one gets a distribution of Chern numbers over the subbands of every Harper band. Now, changing the disorder configuration will give a different distribution of the Chern numbers over the subbands. This reflects the fact that changing the disorder configuration is equivalent to choosing another probability measure on the space of disorder configurations; we already commented that this will change the Chern numbers (see Sec. IV G). However, the sum of all Chern numbers of the subbands is constant and equal to the Chern number of the Harper band, as long as the disorder is weak. As it becomes stronger, the Harper bands overlap energetically; the Hofstadter spectrum loses more of its structure as indicated above. But the distribution of Chern numbers may still be calculated.

Now, for every strength of disorder, averages of the Chern numbers over the disorder configurations can be calculated for every subband. The numerical results then indicate that, on average, the Hall current is essentially carried by states corresponding to one central band situated somewhere at the center of the Harper band. This belief is supported by the calculation of the localization length by the Thouless number method:[90] it diverges at about the energy corresponding to the Hall current carrying subband.[97] In the limit of an infinite sample, one therefore expects the Hall current to be carried by one critical energy, just as in the case without periodic potential. This is, in fact, not very surprising because the semiclassical argument presented in Sec. VI B applies to the present situation as well.

In summary, the numerical calculations suggest the phase diagram[97] shown in an idealized form in Figs. 5 and 6. Starting at rational flux without disorder, the evolution of the current carrying states and their Chern numbers is given as a function of increasing strength of disorder. At some critical values, they amalgamate; the Chern number of the resulting state is then given by

the sum of the merging ones (this sum may also be zero, Hall currents then annihilate each other). To conclude, let us comment on the limiting behavior for strong disorder. In the weak periodic potential approximation, one eventually obtains a Chern number one as for the Landau level. It is an interesting question what happens to the Chern number as the quantum Hall regime breaks down, certainly the corrections to the Kubo formula will be important at very high disorder. In the tight binding approximation, the system evolves to zero Hall current, this reflects the results of Sec. VI A.

VII. INTRODUCTION TO THE FQHE

A. Overview

While in 1982 Thouless, Kohmoto, Nightingale, and den Nijs made an important step towards the understanding of the IQHE,[8] Tsui, Störmer, and Gossard from Bell Laboratories made the surprising discovery that plateaux of the Hall conductivity could be observed not only at integer, but also at fractional multiples of e^2/h.[101] Using a high-quality $GaAs-Al_xGa_{1-x}As$ heterojunction, they observed a plateaux at $\sigma_H/(e^2/h)=\frac{1}{3}$. The corresponding minimum of the direct resistance was also observed. The filling factor was varied by changing the magnetic field while keeping the charge-carrier density fixed. In order to obtain the filling factor $\nu=\frac{1}{3}$, they needed a magnetic field strength of about 15 T. The experiment was undertaken at different values of temperature (0.48, 1.00, 1.65, and 4.15 K) and even at the lowest of them, the accuracy of σ_H on the plateau was much more modest than in the integer effect.

Since this first experiment, the FQHE has been observed for many values of σ_H. The most striking results are the following:

(1) Let $\sigma_H/(e^2/h)=p/q$ with p,q relatively prime integers. Then, q is odd ('odd-denominator rule'). Recently, some deviations of this rule have been observed, but this seems to correspond to non spin-polarized electrons. In the sequence, we will not consider this case.
(2) The observation of the fractional effect requires very clean samples. It is easily destroyed by impurities causing inelastic scattering. Moreover, the stability of the plateaux with respect to temperature and impurity effects is very dissimilar for different fractions. In general, fractions with small denominator seem to be more stable. Thus, the plateaux appear in a well-defined order when lowering the temperature. This gives rise to series of fractions, also called 'hierarchy of states.'
(3) In all devices, the accuracy of the Hall conductivity on the plateaux is much less for fractional than for integer values. Whereas in the IQHE we have $\delta\sigma_H/\sigma_H \approx 10^{-7}$, this is about 10^{-5} at best for the FQHE.

Although we have given a rigorous explanation of the IQHE, the situation is much less satisfactory for the FQHE. The QHE contains two distinct aspects: the existence of plateaux (which, for the IQHE, follows from localization of the states at the Fermi edge) and the integrality (respectively fractionality) of the Hall conductivity on them (which has a topological origin). Usually, these two points are not well separated in the literature on the FQHE. In fact, there are conceptual difficulties in understanding the rôle of localization in the case of interacting electrons. The mechanism which gives rise to the existence of plateaux is quite unclear. The main problem is that no useful Kubo formula is known to calculate the Hall conductivity for interacting Fermions. Instead, one usually follows another strategy: since we know that noninteracting electrons only exhibit an IQHE, it is clear that the FQHE might appear when electron–electron interaction can no longer be neglected. It is generally believed that at certain values of the filling factor, the ground state becomes incompressible owing to the repulsive two-particle interaction. Thus one tries to find explicit ground states which exhibit a fractional Hall conductivity. Then arguments are given to explain that this value remains constant while the filling factor is slightly changed. The most promising states that have been found show a remarkable property: they exhibit particlelike excitations with

some charge $q \neq -e$. The states are only defined for specific values of the filling factor, but it seems that the existence of quasiparticles is stable against perturbations if q/e is a fractional number. Apart from numerical evidence, there is no serious explanation of this property. The stability of quasiparticles is due to an energy gap separating the ground state from the excitations. It is conceptually not clear why the existence of such a gap is related to the fractionality of the quasiparticle charge. No rigorous results are known here.

B. Laughlin's ansatz for the $\nu = 1/m$ ground state

Up to now, most attempts to explain the FQHE—in particular the classical hierarchy of Haldane and Halperin[102,103] as well as Jain's composite-fermions approach[104]—are based on Laughlin's wave functions which describe a finite number of electrons in a finite volume, yielding a filling factor $\nu = 1/m$ with m an odd integer. In the following, we will review the ideas which led Laughlin to his proposal. Following Haldane, we will then construct a class of Hamiltonians for which Laughlin's states are exact ground states.

Before doing this, we want to clarify the rôle of these states and try to justify our interest in them. As we have already mentioned, there does not exist a satisfying theory of the FQHE at the moment. In particular, the universality of the effect and its insensitivity to the exact form of the Hamiltonian is not quite clear. The main advantage of our approach to the IQHE is that we do not need to know the exact form of the one-particle potential, but only have to impose some mild conditions on it. To explain the existence of plateaux and to prove the integrality of $\sigma_H/(e^2/h)$, we did not need to calculate the ground state explicitly. In our opinion, despite the wide use of Laughlin's wave functions in the literature, the explicit form of the ground state does not lie at the heart of the theory, neither for the integer nor for the fractional effect.

Nevertheless, in our theory of the IQHE we were not able to give the explicit values of the Hall conductivity just by means of general considerations. In fact, the value given by the Kubo–Chern formula might depend on the choice of the probability measure P on the hull of the one-particle Hamiltonian. To derive the 'right' values, we computed explicitly the Chern character of the eigenprojections of the Landau Hamiltonian. A homotopy argument then shows that these values remain constant when the disorder is switched on, at least for small values of the potential.

The states which arise from Laughlin's wave functions in the thermodynamical limit might play the same rôle for interacting particles: They yield the explicit values of the conductivity for a Hamiltonian with a restricted two-particle interaction and without one-particle potential. The main task which remains is to show that these values stay constant as the Hamiltonian is deformed continuously. This is going beyond the scope of our article.

We will now come to the description of Laughlin's states. The general Hamiltonian for N electrons is given by

$$H_N = \frac{1}{2m_*} \sum_{j=1}^{N} \left(\mathbf{p}_j + \frac{e}{c} \mathbf{A}(\mathbf{x}_j) \right)^2 + \sum_{1 \leq j < k \leq N} U(|\mathbf{x}_j - \mathbf{x}_k|) + \sum_{j=1}^{N} V(\mathbf{x}_j),$$

where in our case $\nabla \times \mathbf{A}(\mathbf{x}) = \mathscr{B}\mathbf{e}_z$, $\mathscr{B}$=const. Thus we assume that the two-particle interaction depends only on the distance between the particles. Under certain conditions on the functions U and V, H_N is essentially self-adjoint on $C_0^\infty(\mathbf{R}^{2N})$. Furthermore, the N-Fermion space $\wedge^N L^2(\mathbf{R}^2)$ is an invariant subspace. Choosing suitable boundary conditions (we will choose Dirichlet boundary conditions on the disk $\Lambda_R = \{\mathbf{x} \in \mathbf{R}^2 : |\mathbf{x}| \leq R\}$) yields a self-adjoint operator $H_{N,\Lambda}$ in the N-fold tensorproduct $\otimes^N L^2(\Lambda)$. We will denote its restriction to $\wedge^N L^2(\Lambda)$ by the same symbol. Since the Laughlin states should play the same rôle for interacting electrons as the ground state of the Landau–Hamiltonian does for noninteracting ones, we will set $V = 0$. In fact, up to now it is unclear what happens when the one-particle interaction is switched on, even for small values of the potential. The basic idea which leads to Laughlin's states is the following: for

a strong magnetic field, the two-particle interaction can be treated perturbatively; if we start from a system of noninteracting electrons in the lowest Landau level, the excitations to higher Landau levels in the perturbation series can be neglected. Thus, for filling factors $\nu \leqslant 1$, we assume the ground state to be a linear combination of Slater determinants of one-particle wave functions belonging to the lowest Landau level.

For $V=0$, the remaining operator H_N is rotationally invariant. Thus, we can look for joint eigenvectors of energy and total angular momentum. It is clear that, for a repulsive two-particle interaction, the particles will escape to infinity when not restricted to a finite volume. Classically, the trajectory of an electron in a constant magnetic field with angular momentum l and energy E is a circle around the origin with radius $R=\sqrt{l/m_* E}$. Thus, for $E=\hbar\omega_c/2$, restricting the particles to a disk with radius R is essentially equivalent to restricting the angular momentum to values less or equal than $m_*\hbar\omega_c R^2/2 = R^2\hbar/2\ell_B^2 = N\hbar/\nu = N_\Phi\hbar$ (with ℓ_B the magnetic length). Thus, instead of calculating the ground state of $H_{N,\Lambda}$, we are looking for an eigenstate of H_N built of lowest Landau level wave functions with maximal angular momentum N_Φ. Laughlin, strongly influenced by the theory of liquid ^{3}He, proposed the ansatz

$$\psi(z_1,\ldots,z_N)=\prod_{1\leqslant j<k\leqslant N} g(z_k-z_j)$$

for the N-particle ground state. Here and in the following, we set the magnetic length to unit and we use complex coordinates $z=x+\imath y$ in the plane. The lowest Landau level is spanned by the functions $\phi_m \propto z^m e^{-(1/2)|z|^2}$, thus

$$\psi(z_1,\ldots,z_N)=\prod_{1\leqslant j<k\leqslant N} f(z_k-z_j)e^{-(1/2)\sum_{l=1}^N |z_l|^2}$$

with an analytic function f. Since ψ should be anti-symmetric under particle exchange, we have $f(-z)=-f(z)$. Furthermore, as we are looking for eigenstates of total angular momentum M and the operator of angular momentum for one particle is given by $z\partial_z - \bar{z}\partial_{\bar{z}}$, the product in the above formula has to be a polynomial of degree M in each variable z_j. From this, it follows that $f(z)=z^m$ with an odd integer m. We will denote the resulting normalized wave function by $\psi_{N,m}$. The total angular momentum is $M=[N(N-1)/2]m$ and the particles are essentially restricted to a disk with radius $R\approx\sqrt{2m(N-1)}$ (in units of the magnetic length) or, equivalently, $\nu\approx 1/m$.

Before we investigate the question for which Hamiltonians Laughlin's ansatz yields an exact groundstate, we want to make some remarks on the thermodynamical limit. The QHE (the integer as well as the fractional) is assumed to be exact in the case of infinite volume and zero temperature, thus necessarily we have to involve statistical mechanics. Laughlin's states are states for a finite number of particles in a finite volume, i.e., we are working in the canonical ensemble. The Gibbs states are

$$\langle a\rangle_{\beta,N,\Lambda}=\frac{\mathrm{tr}(e^{-\beta H_{N,\Lambda}}a)}{\mathrm{tr}(e^{-\beta H_{N,\Lambda}})},$$

where a is a bounded linear operator in $\wedge^N L^2(\Lambda)$. At zero temperature, they reduce to the vector states

$$\langle a\rangle_{\infty,N,\Lambda}=(\psi_{N,\Lambda}, a\psi_{N,\Lambda}),$$

where $\psi_{N,\Lambda}$ is the (nondegenerate) ground state of $H_{N,\Lambda}$. To investigate the limit $N\to\infty$, $\Lambda\to\mathbf{R}^2$ with fixed particle density $N/|\Lambda|$, we have to include vector states with an arbitrary

number of particles, i.e., we have to work in Fock space. The algebra of observables should be a suitable subalgebra of bounded linear operators in this space. We will choose the canonical anti-commutation relations (CAR) algebra. We recall the definition: Let

$$\mathscr{F}_-(L^2(\mathbf{R}^2)):=\bigoplus_{N=0}^{\infty}\wedge^N L^2(\mathbf{R}^2)$$

denote the Fermi–Fock space. For $\phi\in L^2(\mathbf{R}^2)$ and $\psi\in \oplus\ {}^N L^2(\mathbf{R}^2)$, we shall now define $a^*(\phi)\psi:=\sqrt{N+1}\,\phi\otimes\psi$. Let P_-^N denote the projection onto $\wedge\ {}^N L^2(\mathbf{R}^2)$ in $\otimes^N L^2(\mathbf{R}^2)$, i.e.,

$$P_-^N\psi_1\otimes\ldots\otimes\psi_N=\frac{1}{N!}\sum_{\sigma\in S_N}(-)^\sigma\psi_{\sigma(1)}\otimes\ldots\otimes\psi_{\sigma(N)},$$

where the sum runs over all permutations of $(1,\ldots,N)$. Let $a^*{}_-(\phi)=P_-^{N+1}a^*(\phi)P_-^N$ on $\wedge\ {}^N L^2(\mathbf{R}^2)$, hence, for $\psi\in\wedge^N L^2(\mathbf{R}^2)$

$$a^*{}_-(\phi)\psi(z_1,\ldots,z_{N+1})=\frac{1}{\sqrt{N+1}}\sum_{j=1}^{N+1}(-)^{j+1}\phi(z_j)\psi(z_1,\ldots,\hat{z}_j,\ldots,z_{N+1}),$$

where $\hat{z}_j$ denotes that the jth variable has to be omitted. This yields a bounded linear operator $a_-^*(\phi)$ in $\mathscr{F}_-(L^2(\mathbf{R}^2))$ with $\|a_-^*(\phi)\|=\|\phi\|$. The adjoint operator is given by

$$a_-(\phi)\psi(z_1,\ldots,z_{N-1})=\sqrt{N}\int_{\mathbf{R}^2}d^2z\ \overline{\phi(z)}\psi(z,z_1,\ldots,z_{N-1})$$

for $\psi\in\wedge\ {}^N L^2(\mathbf{R}^2)$. These operators satisfy the CAR

$$\{a_-(\phi),a_-(\psi)\}=\{a_-^*(\phi),a_-^*(\psi)\}=0,\quad \{a_-(\phi),a_-^*(\psi)\}=(\phi,\psi).$$

The CAR algebra over $L^2(\mathbf{R}^2)$, which we will denote by $\mathscr{A}$, is the C^*-algebra of bounded linear operators in $\mathscr{F}_-(L^2(\mathbf{R}^2))$ generated by the identity and the set of operators $\{a_-(\phi):\phi\in L^2(\mathbf{R}^2)\}$.

We regard a state as a linear mapping $\langle\cdot\rangle$ from this algebra into the complex numbers such that $\langle A^*A\rangle\geq 0\ \ \forall A\in\mathscr{A}$ and $\langle 1\rangle=1$. Given a sequence $(\psi_N)_{N\in\mathbf{N}}$ of vectors $\psi_N\in\wedge\ {}^N L^2(\mathbf{R}^2)$, we can ask for the limit of the corresponding vector states. Here, the limit has to be understood in the weak-*-topology, i.e., $\langle\cdot\rangle=\lim_{N\to\infty}\langle\cdot\rangle_N$ if and only if $\langle A\rangle=\lim_{N\to\infty}\langle A\rangle_N\ \ \forall A\in\mathscr{A}$. It is clear that the limit is not a vector state.

As an example, we will treat the case $m=1$. Then, Laughlin's wave function is a single Slater determinant (namely, Vandermonde's determinant). The normalized wave functions are

$$\psi_{N,1}(z_1,\ldots,z_N)=\frac{1}{\sqrt{\pi^N\sqrt{2}^{N(N+1)}1!\ldots N!}}\prod_{1\leq j<k\leq N}(z_k-z_j)e^{-(1/4)\Sigma_{l=1}^N|z_l|^2}.$$

We will denote the limit state by $\langle\cdot\rangle_{\infty,1}$. It is completely determined once the correlation functions $\langle\Pi_{j=1}^m\ a_-^*(\phi_j)\Pi_{k=1}^n\ a_-(\psi_k)\rangle_{\infty,1}$ are known. Obviously, these vanish for $m\neq n$. We will calculate the two-point function explicitly: We have $\langle a_-^*(\phi)a_-(\psi)\rangle_{\infty,1}=(\psi,T\phi)$, where the integral kernel of T is given by

$$(z',Tz)=\lim_{N\to\infty} N\int d^2z_2\cdots d^2z_N\ \overline{\psi_{N,1}(z,z_2,\ldots,z_N)}\psi_{N,1}(z',z_2,\ldots,z_N)$$

$$=\lim_{N\to\infty}\frac{1}{2\pi}\,e^{-1/4(|z|^2+|z'|^2)}\sum_{j=0}^{N-1}\frac{(\bar{z}z'/2)^j}{j!}=\frac{1}{2\pi}\,e^{-(1/4)|z-z'|^2+(i/2)\,z\wedge z'}.$$

Thus, the operator T is just the projection onto the lowest Landau level. As a consequence, we get the correct filling factor $\nu=1$. Furthermore, it can be shown that $\langle\cdot\rangle_{\infty,1}$ is a quasi-free state, i.e., the $2n$-point functions are certain sums over products of two-point functions, for example $\langle a_-^*(\phi_1)a_-^*(\phi_2)a_-(\psi_1)a_-(\psi_2)\rangle_{\infty,1}=\langle a_-^*(\phi_1)a_-(\psi_1)\rangle_{\infty,1}\langle a_-^*(\phi_2)a_-(\psi_2)\rangle_{\infty,1}-\langle a_-^*(\phi_1)a_-(\psi_2)\rangle_{\infty,1}\langle a_-^*(\phi_2)a_-(\psi_1)\rangle_{\infty,1}$; see Ref. 105 for an exact definition. But this is just the limit Gibbs state (for temperature $T=0$ and filling factor $\nu=1$) for noninteracting electrons in a constant magnetic field. Thus in the thermodynamical limit we get the exact ground state.

This can also be seen from the work of Haldane[106] which we will now review. Let Π denote the projection onto the lowest Landau level in the one-particle space. We will investigate the Hamiltonian $H_{N,eff}:=\Pi^{\otimes N}H_N\Pi^{\otimes N}$. The kinetic part just equals $N\hbar\omega/2$. In order to calculate the integral kernel of $\Pi\otimes\Pi U(|z_j-z_k|)\Pi\otimes\Pi$, we perform the unitary transformation

$$\mathcal{U}:L^2(\mathbf{R}^2)\otimes L^2(\mathbf{R}^2)\to L^2(\mathbf{R}^2)\otimes L^2(\mathbf{R}^2),$$

$$(\mathcal{U}\psi)(Z_{jk},z_{jk})=\psi(z_j=Z_{jk}+z_{jk},z_k=Z_{jk}-z_{jk}).$$

Then,

$$\mathcal{U}(\Pi\otimes\Pi)U(|z_j-z_k|)(\Pi\otimes\Pi)\mathcal{U}^{-1}=\Pi\otimes(\Pi U(|z_{jk}|)\Pi).$$

Since $U(|z_{jk}|)$ is rotationally invariant, the second factor in the tensor product is diagonalized by the simultaneous eigenstates of the Landau Hamiltonian and angular momentum. In this basis,

$$(\phi_m,U(|z_{jk}|)\phi_{m'})=U_m\delta_{m,m'}$$

with

$$U_m=\int_0^\infty\frac{dr}{2^{2m+1}m!}\,e^{-(1/4)(r/\ell_{\mathcal{B}})^2}\left(\frac{r}{\ell_{\mathcal{B}}}\right)^{2m+1}U(r),$$

where we have reintroduced the magnetic length $\ell_{\mathcal{B}}$. Thus, in some sense, the projection onto the lowest Landau level yields a 'quantization of interparticle separation.' For example, for the Coulomb interaction $U(r)=e^2/4\pi r$ we have

$$U_m=\frac{e^2}{4\pi}\,\frac{\Gamma(m+1/2)}{2m!}.$$

For $m\to\infty$, this falls off as $1/\sqrt{m}$. It is remarkable that the coefficients U_m are independent of $\ell_{\mathcal{B}}$ exactly for the Coulomb interaction.

As a result, we state: The set of functions

$$\psi(Z,z)\propto Z^M z^m e^{-(1/4)(|z|^2+|Z|^2)}\propto(z_j+z_k)^M(z_j-z_k)^m e^{-(1/4)(|z_j|^2+|z_k|^2)}$$

with $M,m\in\{0,1,2,\ldots\}$ forms a basis of eigenfunctions of $\Pi\otimes\Pi U(|z_j-z_k|)\Pi\otimes\Pi$. Each eigenvalue U_m is infinitely degenerate (according to an arbitrary angular momentum of the center-of mass motion).

The general two-particle wave function in the lowest Landau level with angular momentum maximal equal to N_Φ is given by

$$\psi(z_1,z_2)=P_{N_\Phi}(z_1,z_2)e^{-(1/4)(|z_1|^2+|z_2|^2)}$$

where P_{N_Φ} is a polynomial of degree N_Φ in each of its variables. These states form a $\binom{N_\Phi}{2}$-dimensional subspace, but this subspace is not invariant under the two-particle interaction. Let us assume that the sequence U_m decreases monotonically. Then the desired ground state would be a state with relative angular momentum N_Φ, or, since we are dealing with Fermions and therefore the relative angular momentum has to be odd, $N_\Phi-1$. In the latter case, we have a degenerate ground state since we can choose $M=0$ or $M=1$. Let us remark that these functions are not equal to Laughlin's ansatz for $N=2$.

In the case of more than two particles, it is in general not possible to give the ground state. The general N-particle wave function in the lowest Landau level with maximal angular momentum equal to N_Φ is given by

$$\psi(z_1,...,z_N)=P_{N_\Phi}(z_1,...,z_N)e^{-(1/4)\sum_{l=1}^{N}|z_l|^2},$$

where again $P_{N_\Phi}$is a polynomial of degree N_Φ in each of its variables. These states form a $\binom{N_\Phi+1}{N}$-dimensional subspace. Let us require that only pairs with a relative angular momentum of at least m_0 appear in the wave function. This is possible if and only if $m_0\leqslant N_\Phi/(N-1)\to 1/\nu$, and the general form of the wave function is given by

$$\psi(z_1,...,z_N)=P_{N_\Phi-m_0(N-1)}\prod_{1\leqslant j<k\leqslant N}(z_k-z_j)^{m_0}e^{-(1/4)\sum_{l=1}^{N}|z_l|^2}.$$

Let us now consider the truncated Hamiltonian which results from setting $U_m=0$ for $m\geqslant m_0$. Then, the above ψ is an eigenstate with energy 0. Furthermore, if we assume $U_m=0$ for $m<m_0$, it is a ground state. In general, for $\nu<1/m_0$, the ground state is degenerate due to the polynomial prefactor $P_{N_\Phi-m_0(N-1)}$. Let us assume that m_0 is odd. Then, to get a Fermionic state, this prefactor has to be symmetric under particle exchange. Thus, if we assume $N_\Phi-m_0(N-1)=0$ which is equivalent to $1/m_0=\nu-1/N_\Phi\to\nu$, we get a unique ground state which is equal to Laughlin's state $\psi_{N,m}$ with $m=m_0$. For m_0 even, $P_{N_\Phi-m_0(N-1)}$ has to be antisymmetric. There is (up to a constant factor) a unique possibility if the polynomial is of degree one, namely $P_1(z_1,...,z_N)=\Pi_{1\leqslant j<k\leqslant N}(z_k-z_j)$. In this case, we get $\psi_{N,m}$ as a unique ground state for $\nu=1/m$ and $m=m_0+1$.

Let us recapitulate the result: if we project the Hamiltonian onto the lowest Landau level and assume that all coefficients U_m vanish for m greater or equal to some m_0 and $U_m>0$ otherwise, then the unique (infinite volume) ground state with filling factor $\nu=1/m_0$ for m_0 odd respectively $\nu=1/m_0+1$ for m_0 even is the limit state which arises from Laughlin's wave functions ψ_{N,m_0} or ψ_{N,m_0+1} respectively.

In fact, we have not proved this rigorously. First, it remains to show that Laughlin's wave functions indeed yield the correct filling factor in the thermodynamic limit. This has been shown by Laughlin by means of a somewhat heuristic argument.[107] Then, it has to be shown that the restriction of angular momentum indeed yields the same thermodynamic limit as a classical boundary condition. This should not be too hard to prove.

C. The elementary theory of the $\nu=1/m$-FQHE

The elementary theory of the $\nu=(1/m)$-FQHE is solely based on Laughlin's states. Thus we assume in the following that the correct ground state is indeed given by their thermodynamical limit. We then have to explain that

(1) the Hall conductivity for the state given by $\psi_{N,m}$ equals e^2/mh,
(2) plateaux of the Hall conductivity arise when the one-particle potential is switched on.

We remark that the first point is just the classical result $\sigma_H/(e^2/h)=\nu$. Indeed, for $V=0$, the classical result should hold for any value of the filling factor.

To calculate σ_H, Laughlin[107] proposed a Gedanken experiment similar to that we have already described in Sec. II E. We consider the N-particle wave function $\psi_{N,m}$. We introduce an infinitely thin solenoid at the origin, perpendicular to the disk in which the particles are confined. Then we force adiabatically one flux quantum $\Phi_0=hc/e$ through the flux tube. We assume that $\psi_{N,m}$ is the nondegenerate ground state for some Hamiltonian and that there is a gap between the ground state energy and the rest of the spectrum. Let us mention that this 'gap-condition' is a sufficient condition for our considerations, but in general need not be necessary for $\sigma_H/(e^2/h)$ to be equal to $1/m$. After the flux quantum has been added, the Hamiltonian has not changed up to a gauge transformation. Since we can apply the adiabatic theorem, we know that the system is still in an eigenstate of the Hamiltonian. The single-particle wave functions $z^m e^{-(1/4)|z|^2}$ evolve, up to a phase factor and normalization, to $z^{m+1}e^{-(1/4)|z|^2}$ during this operation. The particles which were essentially confined in a disk with radius $R=\sqrt{2N_\Phi}$ at the beginning are now restricted to the radius $R'=\sqrt{2(N_\Phi+1)}$. These two disks differ by an area 2π, and since the charge density equals uniformly $e/2\pi m$, a charge e/m has left the original disk while adding the flux quantum. Similar to the IQHE, this 'gauge argument' does not prove directly the fractionality of the Hall conductivity, but shows that the classical result $\sigma_H/(e^2/h)=\nu$ is valid. Fractionality of $\sigma_H/(e^2/h)$ follows from the fractionality of the filling factor.

Let us discuss the excited state which appears in Laughlin's Gedanken experiment in the case of Haldane's truncated Hamiltonian. After the flux quantum has been added, the single-particle angular momentum is restricted to values less than or equal to $N_\Phi+1$. Since we are looking for an excited state, the relative angular momentum for each pair is still at least m. Thus,

$$\psi(z_1,\ldots,z_N)=P_1(z_1,\ldots,z_N)\psi_{N,m}(z_1,\ldots,z_N).$$

In general, $P_1(z_1,\ldots,z_N)\propto\Pi_{j=1}^N(z_j-u_j)$ with arbitrary u_j. But as P_1 has to be symmetric under particle exchange, all u_j have to be equal. This yields the excited states already proposed by Laughlin:[107]

$$S_{z_0}\psi_{N,m}(z_1,\ldots,z_N)=\prod_{j=1}^N(z_j-z_0)\psi_{N,m}(z_1,\ldots,z_N).$$

The adjoint operator is given by

$$S_{z_0}^*\psi_{N,m}(z_1,\ldots,z_N)=e^{-(1/4)\sum_{l=1}^N|z_l|^2}\prod_{j=1}^N(2\partial_{z_j}-\overline{z_0})\prod_{1\le k<l\le N}(z_l-z_k)^m.$$

The operators S_{z_0} and $S_{z_0}^*$ are called 'quasiparticle' respectively 'quasihole' creation operators for the following reasons: Denote by $\Phi_{N,m,z}$ the N-particle state $e^{-(1/4m)|z|^2}S_z\psi_{N,m}$. Then, one can show that

$$(\Phi_{N,m,z},\Phi_{N,m,z'})\propto e^{-(1/4m)(|z|^2+|z'|^2-2\bar{z}z')}.$$

Thus $\Phi_{z,m}$ looks like the wave function of a 'quasiparticle' with charge e/m in the lowest Landau level. More formally: Let $\psi \in L_2(\mathbf{R}^2)$ (a quasiparticle wave function). Define the N-electron state

$$Q_{N,m}\psi := \frac{1}{\sqrt{2\pi m}} \int d^2\eta\ \overline{\psi(\eta)} e^{-(1/4m)|\eta|^2} S_\eta \psi_{N,m}.$$

Then, for $\psi_z(\eta)=(1/\sqrt{2\pi m})e^{-(1/4m)(|z|^2+|\eta|^2-2\bar{z}\eta)}$, we have $Q_{N,m}\psi_z=\Phi_{N,m,z}$ and moreover $(\psi_z,\psi_{z'})=(Q_{N,m}\psi_z,Q_{N,m}\psi_{z'})$.

Of course, the state $\Phi_{z,m}$ is in general not an exact eigenstate of the truncated N-particle Hamiltonian. But as $N\to\infty$, its overlap with an eigenstate becomes larger and larger. Let us assume, for simplicity, that it is an eigenstate. Since the quasiparticle state ψ_z is an eigenstate of the Hamiltonian

$$H_q=\left(\frac{1}{\iota}\nabla_\eta+\frac{e}{m}\mathbf{A}(\eta)\right)^2,$$

we have

$$(\Phi_{N,m,z},H_{N,eff}\Phi_{N,m,z})\propto(\psi_z,H_q\psi_z),$$

where, for simplicity, we have denoted also the truncated Hamiltonian by $H_{N,eff}$. One is led to define a quasiparticle potential V_q by the equation

$$\left(\Phi_{N,m,z},\sum_{j=1}^{N} V(z_j)\Phi_{N,m,z}\right)\propto(\psi_z,V_q\psi_z)=\frac{1}{2\pi m}\int d^2\eta\ e^{-(1/2m)|z-\eta|^2}V_q(\eta).$$

Thus, we have defined an effective Hamiltonian which governs the quasiparticle motion. This yields a qualitative explanation for the existence of plateaux: If the quasiparticles which are created by an electric field are localized due to randomness of the potential V_q, they do not contribute to the Hall conductivity. The FQHE looks like the IQHE for quasiparticles.

But this only concerns the mechanism which causes the existence of plateaux. The fractionality of $\sigma_H/(e^2/h)$ is due only to the fact that the charge density of Laughlin's states is exactly $e/2\pi m$. As we have shown, these states are indeed the nondegenerate ground states for some Hamiltonian, but it is not clear at all why the quasiparticle charge should be invariant under slight deformations of the ground state. Thus, the topological origin of the FQHE remains uncovered.

D. The rôle of gauge invariance and incompressibility

We have tried to emphasize that the main problem of the common approach to the FQHE is the lack of a Kubo formula to calculate the Hall conductivity. To calculate it for Laughlin's states, we used gauge invariance and assumed the existence of a gap separating the ground state energy from the rest of the spectrum. These two assumptions—gauge invariance and incompressibility—also form requirements for a completely different approach which is based on a work of Halperin[51] and has been developed further, e.g., by Wen[108] and Fröhlich *et al.*[109–114] We will particularly refer to the work of the latter group.

Let us first look at incompressibility from a classical point of view. How does the system respond to external "small" fields $\mathscr{E}$ and $\mathscr{B}$ (where the latter does *not* include the strong magnetic field which causes the quantum Hall state)? In $(2+1)$-dimensional space–time (we will only consider cases where space is a subset $\Lambda\subset\mathbf{R}^2$—in particular a disk around the origin—with the

usual 'flat' metric and space-time will then be $\mathbf{R}\times\Lambda$), we only have two components of the electric and one component of the magnetic field. In covariant notation, we have the electromagnetic-field tensor $F=\frac{1}{2}F_{\mu\nu}\,dx^{\mu}\wedge dx^{\nu}$ with $x^0=ct$ and

$$(F_{\mu\nu})_{\mu,\nu\in\{0,1,2\}}=\begin{pmatrix} 0 & \mathscr{E}_x & \mathscr{E}_y \\ -\mathscr{E}_x & 0 & -\mathscr{B}_z \\ -\mathscr{E}_y & \mathscr{B}_z & 0 \end{pmatrix}.$$

The homogeneous Maxwell equations $dF=0$ reduce to Faraday's induction law

$$\frac{1}{c}\,\partial_t\mathscr{B}_z+\nabla\wedge\mathscr{E}=0.$$

Let $\mathbf{j}$ denote the electric current density. Then we have the Ohm–Hall law (respectively, the definition of the conductivity tensor σ) $\mathbf{j}=\sigma\mathscr{E}$. Furthermore, $\sigma=\rho^{-1}$, where ρ is the symmetric resistivity tensor. Incompressibility means that the diagonal components of ρ vanish and hence

$$\sigma=\begin{pmatrix} 0 & \sigma_H \\ -\sigma_H & 0 \end{pmatrix}.$$

Furthermore, we have the continuity equation (i.e., charge conservation)

$$\frac{1}{c}\,\partial_t j^0+\nabla\mathbf{j}=0,$$

where cj^0 denotes charge density minus the uniform charge density of the unperturbed quantum Hall state. From these three basic equations, we get $\partial_t j^0=\sigma_H\partial_t\mathscr{B}$. Since, with our above definition of j^0, we have $j^0=0$ for $\mathscr{B}=0$, the integration of this equation yields $j^0=\sigma_H\mathscr{B}$.

We introduce the 2-form J dual to $(j^0,\mathbf{j})$:

$$J=\frac{1}{2}J_{\mu\nu}\,dx^{\mu}\wedge dx^{\nu},\quad (J_{\mu\nu})_{\mu,\nu\in\{0,1,2\}}=\begin{pmatrix} 0 & j_y & -j_x \\ -j_y & 0 & j^0 \\ j_x & -j^0 & 0 \end{pmatrix}.$$

Then, the continuity equation reads $dJ=0$ and we can combine the Ohm–Hall law and the connection between j^0 and $\mathscr{B}_z$ into one equation: $J=-\sigma_H F$. Since we work in simply connected space–time, we have

$$J=db,\quad F=da$$

with 1-forms a and b. Hence, $db=-\sigma_H da$. This is the Euler–Lagrange equation for some action. Whether we regard it as an equation for a or for b is just a matter of taste, corresponding to the choice that we can either create a current by applying an external field or, equivalently, can create a Hall voltage by forcing some current. For our purpose, we will choose the latter possibility. Thus, our 'dynamical variable' is the gauge potential whereas we keep the potential for the current density fixed. A possible choice for the action is

$$S_{\Lambda,b}(a)=\frac{\sigma_H}{2}\int_{\Lambda\times\mathbf{R}} a\wedge da+\int_{\Lambda\times\mathbf{R}} b\wedge da+\text{b.t.},$$

where, since we work in a finite volume, we have to include some (unknown) boundary term 'b.t.'

Can these phenomenological equations be derived from a microscopical description? And is there a connection between the quantization of the classical theory (by means of path integrals) and our original quantum mechanical problem? Formally, the Hamiltonian $H_\Lambda = \oplus_{N=0}^{\infty} H_{N,\Lambda}$ (with $H_{N,\Lambda}$ as in the previous sections) can be described by means of 'fields' Ψ and Ψ^* which satisfy at each time t the CAR

$$\{\Psi(\mathbf{x},t),\Psi(\mathbf{y},t)\}=\{\Psi^*(\mathbf{x},t),\Psi^*(\mathbf{y},t)\}=0, \quad \{\Psi(\mathbf{x},t),\Psi^*(\mathbf{y},t)\}=\delta(\mathbf{x}-\mathbf{y}).$$

The fields act on the Fermi–Fock space $\mathscr{F}_-(L_2(\mathbf{R}^2))$. The particle-number operator is given by $N_\Lambda=\int_\Lambda \Psi^*(\mathbf{x},t)\Psi(\mathbf{x},t)$. We will work in the grand-canonical ensemble. We have

$$\begin{aligned} H_\Lambda(t)-\mu N_\Lambda = &-\frac{\hbar^2}{2m_*}\int_\Lambda d^2x \left[\nabla-\frac{\imath e}{\hbar c}\mathbf{a}_{tot}(\mathbf{x},t)\right]\Psi^*(\mathbf{x},t)\left[\nabla+\frac{\imath e}{\hbar c}\mathbf{a}_{tot}(\mathbf{x},t)\right]\Psi(\mathbf{x},t) \\ &+\int_\Lambda d^2x\ \Psi^*(\mathbf{x},t)[V(\mathbf{x})-\mu+ea^0]\Psi(\mathbf{x},t)+ \\ &+\frac{1}{2}\int_{\Lambda\times\Lambda} d^2x d^2y\ \Psi^*(\mathbf{x},t)\Psi^*(\mathbf{y},t)U(|\mathbf{x}-\mathbf{y}|)\Psi(\mathbf{y},t)\Psi(\mathbf{x},t). \end{aligned}$$

Here, $\mathbf{a}_{tot}$ represents the potential for the total magnetic field (i.e., the potential $\mathbf{A}$ for the constant field $\mathscr{B}\mathbf{e}_z$ plus some—in general time-dependent—source term $\mathbf{a}$) and a^0 also is a source term.

We then have the Heisenberg equation of motion

$$\partial_t\Psi(\mathbf{x},t)=\frac{\imath}{\hbar}[H_\Lambda(t)-\mu N_\Lambda,\Psi(\mathbf{x},t)].$$

But this is the Euler–Lagrange equation derived from the action

$$S_\Lambda(\Psi^*,\Psi)=\int_{\mathbf{R}} dt\left\{\int_\Lambda d^2x\ \Psi^*(\mathbf{x},t)\imath\hbar\partial_t\Psi(\mathbf{x},t)-H_\Lambda(t)+\mu N_\Lambda\right\}.$$

The partition function (at zero temperature and chemical potential μ) is given by the path integral

$$Z_\Lambda=\int \mathscr{D}\Psi^*\ \mathscr{D}\Psi e^{(i/\hbar)S_\Lambda(\Psi^*,\Psi)}.$$

We define the effective action as $S_\Lambda^{eff}:=(\hbar/\imath)\ln Z_\Lambda$. Regarded as a functional of the gauge potential $(a_{tot})_\mu\, dx^\mu=a^0\, dx^0-\mathbf{a}_{tot}\, d\mathbf{x}$, it is the generating functional for connected time-ordered Green functions of the current-density operators

$$j^0=\Psi^*\Psi, \quad \mathbf{j}=-\frac{\imath\hbar}{2m_*c}\left\{\left[\left(\nabla-\frac{\imath e}{\hbar c}\,\mathbf{a}_{tot}\right)\Psi^*\right]\Psi-\Psi^*\left(\nabla+\frac{\imath e}{\hbar c}\,\mathbf{a}_{tot}\right)\Psi\right\}.$$

We have

$$\langle Tj^{\mu_1}(x_1)\cdots j^{\mu_n}(x_n)\rangle^c=\imath^{(n+1)}\frac{\delta}{\delta a_{\mu_1}(x_1)}\cdots\frac{\delta}{\delta a_{\mu_n}(x_n)}\,S_\Lambda^{eff}(a^0,\mathbf{a}).$$

We now perform a scale transformation, which means the following: We enlarge the domain Λ by multiplying the space–time $\mathbf{R}\times\Lambda$ with a scalar $\lambda>1$. We introduce new coordinates $\xi=(\xi^0,\boldsymbol{\xi})=\lambda^{-1}(x^0,\mathbf{x})$ (thus $\boldsymbol{\xi}\in\Lambda$) and define the rescaled sources $a_\mu^\lambda(\lambda\xi):=\lambda^{-1}a_\mu(\xi)$, such

that $a_\mu^\lambda(\lambda\xi)d(\lambda\xi)=a_\mu(\xi)d\xi$. The 'scaling-limit' $\lambda\to\infty$ corresponds to a thermodynamical, but also adiabatic, limit. We assume some 'strong-clustering property', namely, we require that for $n\leqslant 3$ the distributions

$$\phi^{\mu_1\cdots\mu_n}(\xi_1,\ldots,\xi_n):=\frac{\imath^{n+1}}{n!}\lim_{\lambda\to\infty}\lambda^{2n}\langle Tj^{\mu_1}(\lambda\xi_1)\cdots j^{\mu_n}(\lambda\xi_n)\rangle^c$$

are *local* distributions, i.e.,

$$\operatorname{supp}\phi^{\mu_1\cdots\mu_n}=\{(\xi_1,\ldots,\xi_n)\in(\mathbf{R}\times\Lambda)^n:\xi_1=\cdots=\xi_n\}.$$

Then, assuming some differentiability condition on S_Λ^{eff}, we have up to second order (with $A=-\mathbf{A}d\mathbf{x}$):

$$S_{\lambda\Lambda}^{eff}(A+a^\lambda)=S_{\lambda\Lambda}^{eff}(A)+\int_{\lambda(\mathbf{R}\times\Lambda)}d^3x\ \frac{\delta}{\delta a_\mu(x)}\bigg|_{a=A}S_{\lambda\Lambda}^{eff}(a)\ a_\mu^\lambda(x)$$
$$+\int_{[\lambda(\mathbf{R}\times\Lambda)]^2}d^3xd^3y\ \frac{\delta^2}{\delta a_\mu(x)\delta a_\nu(y)}\bigg|_{a=A}S_{\lambda\Lambda}^{eff}(a)\ a_\mu^\lambda(x)a_\nu^\lambda(y).$$

Now, $\delta/\delta a_\mu(x)|_{a=A}S_{\lambda\Lambda}^{eff}(a)=-\langle j^\mu(x)\rangle^c|_{a=A}$. We define $j_c^\mu(\xi):=\lim_{\lambda\to\infty}\lambda^2\langle j^\mu(\lambda\xi)\rangle^c|_{a=A}$ and J_c as the 2-form dual to j_c.

Furthermore, we have

$$\frac{1}{2}\frac{\delta^2}{\delta a_\mu(x)\delta a_\nu(y)}\bigg|_{a=A}S_{\lambda\Lambda}^{eff}(a)=-\frac{\imath}{2}\langle j_\mu(x)j^\nu(y)\rangle^c\bigg|_{a=A}\overset{\lambda\to\infty}{\to}\phi^{\mu\nu}(\lambda^{-1}x,\lambda^{-1}y).$$

By the above clustering property, we can split $\phi^{\mu\nu}$ in the manner

$$\phi^{\mu\nu}(\xi,\eta)=\alpha\varepsilon^{\mu\nu\rho}(\partial_\rho\delta)(\xi-\eta)+R^{\mu\nu}(\xi,\eta),$$

where α is a constant and $R^{\mu\nu}$ consists of second or higher derivatives of the δ-function.

It can be shown[112] that neither the part with $R^{\mu\nu}$ in the second-order term nor any of the terms of higher order contribute to the scaling limit up to some boundary term. In the language of field theory, they are irrelevant, whereas the first-order term is relevant and the remaining part of second order (the 'Chern–Simons term') is marginal. Hence, a simple substitution in the remaining integrals leads to

$$S_{\lambda\Lambda}^{eff}(A+a)-S_{\lambda\Lambda}^{eff}(A)\overset{\lambda\infty\to}{\to}\int_{\Lambda\times\mathbf{R}}J_c\wedge a+\alpha\int_{\Lambda\times\mathbf{R}}a\wedge da+\text{b.t.}$$

This is exactly our phenomenological action; thus the clustering conditions indeed correspond to incompressibility. By comparison, we expect that $\alpha=\sigma_H/2$. Let us check this explicitly. We reintroduce the non-rescaled variables:

$$S_{\lambda\Lambda}^{eff}(A+a)\overset{\lambda\to\infty}{\to}S_{\lambda\Lambda}^{eff}(A)-\int_{\lambda(\Lambda\times\mathbf{R})}d^3x\ \langle j^\mu(x)\rangle^c\bigg|_A a_\mu^\lambda(x)$$
$$+\alpha\int_{\lambda(\Lambda\times\mathbf{R})}d^3x\ \varepsilon^{\mu\nu\rho}a_m^\lambda u(x)\partial_\nu a_\rho^\lambda(x).$$

By calculating the functional derivative, we obtain

$$\langle j^\mu(x)\rangle^c\Big|_{A+a^\lambda} = \langle j^\mu(x)\rangle^c\Big|_A - 2\varepsilon^{\mu\nu\rho}\partial_\nu a^\lambda_\rho(x).$$

For example, let us take as source the potential for a constant electric field in y direction: $a^\lambda_0(x^0,\mathbf{x}) = -\mathscr{E}x_2$, $\mathbf{a}^\lambda = \mathbf{0}$. Then $\langle j^1(x)\rangle^c|_{A+a^\lambda} = \langle j^1(x)\rangle^c|_A - 2\alpha\mathscr{E}$. Thus indeed we have $\alpha = \sigma_H/2$.

The boundary term in the effective action is in general unknown, but we have the requirement that the whole action is gauge-invariant, i.e., $S^{eff}_\Lambda(a+d\chi) = S^{eff}_\Lambda(a)$ for arbitrary smooth functions χ. From this, it follows that

$$\text{b.t. } (a+d\chi) - \text{b.t. } (a) = \frac{\sigma_H}{2}\int_{\partial\Lambda\times\mathbf{R}} d\chi\wedge a.$$

We introduce light-cone coordinates $u_\pm = (1/\sqrt{2})(vt \pm (\theta/2\pi)L)$ on $\partial\Lambda\times\mathbf{R}$, where v is some (*a priori* arbitrary) velocity, θ is the angle of polar coordinates in the plane and L is the circumference of the disc Λ. Then the above functional equation for the boundary term has the general solution

$$\text{b.t.}(a) = \int_{\partial\Lambda} d^2u \left(a_+a_- - 2a_+\frac{\partial_-^2}{\Box}a_- \right) + W(a) =: \Delta(a) + W(a),$$

where $a|_{\partial\Lambda} = a_+du_+ + a_-du_-$, $\partial_\pm = \partial/\partial u_\pm$, $\Box = 2\partial_+\partial_-$ and W is some arbitrary gauge-invariant function.

Δ is known as the generating functional for time-ordered connected Green functions of chiral $U(1)$-currents. We just state the result and refer the reader to Refs. 115 and 116: Assume that there are current-density operators in two-dimensional space–time whose time-ordered connected Green functions are given by the functional derivatives of Δ. These operators, then are sums of derivatives of n massless scalar fields $\phi_1,...,\phi_n$ for some $n\in\{1,2,...\}$. The dynamics of these fields is determined by the action

$$S_a(\phi_1,...,\phi_n) = \frac{1}{2}\int_{\partial\Lambda} d^2u\ \partial_+\phi^t K\partial_-\phi$$
$$- \int_{\partial\Lambda} d^2u\ [a^t_-\partial_+\phi - (\partial_-\phi - K^{-1}a_-)^t a_+] + \frac{\kappa}{2}\int_{\partial\Lambda} d^2u\ a_-a_+,$$

where K is a positive-definite $n\times n$-matrix, $\phi^t = (\phi_1,...,\phi_n)$, $a^t = (a,...,a)$, and $a^t_\pm = (a_\pm,...,a_\pm)$, $\kappa = \Sigma^n_{j,k=1}\,(K^{-1})_{jk}$ and we have to impose the chirality constraint

$$\partial_-\phi - K^{-1}a_- = 0.$$

Indeed it can be checked that the effective action of this theory, which is defined analogously to S^{eff}_Λ, is given by $(\kappa/2)\Delta$. Thus after performing the 'Abelian Bosonization,' we see that the total effective action is gauge-invariant if and only if $\sigma_H = \kappa$.

Up to now, we have (apart from positive-definiteness) no requirements on the matrix K and thus $\sigma_H/(e^2/h)$ could be any real number. What are the requirements for a rational value of the Hall conductivity? It turns out that we have to impose some *further* condition. We will sketch the idea and omit the details. For simplicity, we will restrict ourselves to the case $n=1$. Thus, we have a single massless scalar field. We introduce the variables $\hat{a}_\pm := (e/\hbar)\,a_\pm$ (which have the dimension of an inverse length) and $\hat{K} := (e^2/h)K$ (which is dimensionless) and we normalize the field ϕ such that the action is given by

$$S_{\hat{a}}(\phi)=\frac{1}{2\pi}\int_{\mathbf{R}\times\partial\Lambda} d^2u \left\{\frac{\hat{K}}{2}(\partial_+\phi)(\partial_-\phi)-\left[\hat{a}_-\partial_+\phi-\left(\partial_-\phi-\frac{1}{\hat{K}}\hat{a}_-\right)\hat{a}_+\right]+\frac{1}{2\hat{K}}\hat{a}_-\hat{a}_+\right\}.$$

We can break down ϕ into its chiral components:

$$\phi(u_+,u_-)=\phi_L(u_+)+\phi_R(u_-).$$

'L' and 'R' correspond to the sign of the charge of the (quasi-)particles described by ϕ. Let us treat the part ϕ_L. We have the current-density operator $J_L(u_+)\propto\partial_+\phi_L(u_+)$, where at the moment it is not quite clear how to define the constant of proportionality. We will choose it such that the corresponding charge operator is given by

$$Q_L=-\frac{e}{2\pi}\oint d\theta\left(\partial_+\phi_L-\frac{1}{\hat{K}}\hat{a}_+\right).$$

The form of Q_L is clear for $\hat{a}=0$ and the general case follows from gauge invariance. Of course, this determines Q_L only up to some constant factor. We will come back to this point which seems to be crucial to us. We define 'vertex operators'

$$V_n(u_+)=:e^{in\phi_L(u_+)}:,$$

where ': :' denotes normal ordering. These operators obey the important relations

$$V_m(u_+)V_n(v_+)=e^{\pm i\pi(mn/\hat{K})}V_n(v_+)V_m(u_+),$$

$$[Q_L,\ V_n(u_+)]=-\frac{ne}{\hat{K}}V_n(u_+).$$

Thus, $V_n(u_+)$ creates a quasiparticle of 'charge' $Q_L=-ne/\hat{K}$. If we create two such quasiparticles and interchange them, we get a phase factor $e^{\pm i\pi(n^2/\hat{K})}$, i.e., these quasiparticles obey *fractional statistics* in general. If $\hat{K}=n$, then we get a phase factor $e^{\pm i\pi\hat{K}}$. We now come to the announced additional assumption: Assume that the quasiparticles with 'charge' $-e$ created by vertex operators are electrons and therefore obey Fermi statistics. Then we get the celebrated result

$$\hat{K}=2l+1\Leftrightarrow\sigma_H=\frac{1}{2l+1}\frac{e^2}{h},\quad l\in\mathbf{Z}.$$

Let us just mention that in the general case, i.e. with more than one field, $\sigma_H/(e^2/h)$ is still a rational number, namely,

$$\sigma_H=\sum_{j,k=1}^{N}(K^{-1})_{jk}\frac{e^2}{h},\quad K_{jk}\in\mathbf{Z},\quad K_{jj}\in 2\mathbf{Z}+1.$$

Furthermore, if we also take quasiparticles with positive charge into account, $\sigma_H/(e^2/h)$ is the difference between two such rational numbers.

We end this section with two remarks: First, it seems to us that the last assumption is more or less ad hoc. Why should we exclude quasiparticles with fractional statistics and charge $-e$? Even if we assume that the electron is among the possible excitations, it is not clear that we can create this excitation by just applying one vertex operator.

Second, there is still the question of normalization of the charge operator. Assume that we change its normalization by multiplying it with a factor γ. Then, the vortex operator $V_n(u_+)$ creates a quasiparticle with 'charge' $-\gamma ne/\hat{K}$. Interchanging two such quasiparticles with 'charge' $-e$ would then yield a phase factor $e^{\pm i\pi(\hat{K}/\gamma^2)}$. Thus, by varying γ, $\hat{K}$ could be any real number! So what is the reason for the above normalization of Q_L? From its definition and Stokes' theorem, we see immediately that adding one flux quantum $\phi_0 = hc/e$ creates a 'charge' $-e/\hat{K}$ at the boundary. On the other hand, we know from Laughlin's gauge argument that a charge $-e\nu$ leaves the bulk when the flux quantum is added (and, in this framework, then circles along the boundary). But this means that $\hat{K} = \nu$, or $\sigma_H = \nu(e^2/h)$, which is the *classical* result! Our feeling is that we find the same problem as in our discussion of Laughlin's argument in Sec. II E. Without localization, the classical result $\sigma_H = \nu(e^2/h)$ should hold for any ν. When will the value of σ_H be robust against turning on some disorder? Presumably, this is exactly the case if ν belongs to the set of fractions which yield an incompressible state. But no serious arguments are known for this hypothesis, apart from the fact that it fits well with numerical and experimental results. Thus to explain the existence of plateaux in the FQHE is still an open problem, but we believe that the reviewed results form an important step to its solution.

ACKNOWLEDGMENTS

This paper has benefited from many contacts during the last ten years. It is almost impossible to give the list of all colleagues who contributed to these discussions. One of us (J.B.) gives his special thanks to Y. Avron, R. Seiler, B. Simon, and, of course, to A. Connes for his outstanding contribution, continuous support, and warm encouragements. We would also like to thank T. Ziman and Y. Tan for their help during the writing of this paper. J.B. had very helpful discussions with B.I. Shklovskii about the Mott conductiviy.

Two of us (J.B. and H.S.B.) would like to thank the FORTH and the University of Crete for providing support during the final edition of this paper.

[a]E-mail: jeanbel@siberia.ups-tlse.fr
[b]E-mail: andreas@tolosa.ups-tlse.fr
[c]E-mail: hermann@tolosa.ups-tlse.fr
[d]URA 505, CNRS and Laboratoire de Physique Quantique, 118, Route de Narbonne, 31062-Toulouse Cedex, France.

[1] E. Hall, Am. J. Math. **2**, 287 (1879); and in *Quantum Hall effect: a perspective*, edited by A. Mac Donald (Kluwer Academic, Dordrecht, 1989).
[2] K. v. Klitzing, G. Dorda, and M. Pepper, Phys. Rev. Lett. **45**, 494 (1980).
[3] M. Stone Ed., *The Quantum Hall Effect* (World Scientific, Singapore, 1992).
[4] A. Connes, Publ. IHES **62**, 257 (1986).
[5] A. Connes, *Géométrie non commutative* (InterEditons, Paris, 1990).
[6] A. Connes, Noncommutative Geometry, preprint IHES m/93/54 (1993).
[7] R.B. Laughlin, Phys. Rev. B **23**, 5632 (1981).
[8] D. Thouless, M. Kohmoto, M. Nightingale, and M. den Nijs, Phys. Rev. Lett. **49**, 405 (1982).
[9] R.E. Prange, Phys. Rev. B **23**, 4802 (1981).
[10] R. Joynt and R. Prange, Phys. Rev. B **29**, 3303 (1984).
[11] D.J. Thousless, J. Phys. C **14**, 3475 (1981).
[12] J.E. Avron, R. Seiler, and B. Simon, Phys. Rev. Lett. **51**, 51 (1983).
[13] H. Kunz, Commun. Math. Phys. **112**, 121 (1987).
[14] J. Bellissard, in *Statistical Mechanics and Field Theory: Mathematical Aspects*, *Lecture Notes in Physics* **257**, edited by T. Dorlas, M. Hugenholtz, and M. Winnink (Springer-Verlag, Berlin, 1986).
[15] J. Bellissard, in *Operator algebras and applications*, Vol. 2, edited by E. Evans and M. Takesaki (Cambridge University, Cambridge, 1988).
[16] J. Bellissard, in *Proceedings of the Bad Schandau conference on localization*, edited by Ziesche & Weller (Teubner-Verlag, Leipzig, 1987).
[17] J.E. Avron, R. Seiler, and B. Simon, Commun. Math. Phys. **159**, 399 (1994).
[18] N.W. Ashcroft, and N.D. Mermin, *Solid State Physics* (Saunders, Philadelphia, 1976).
[19] T. Ando, A. Fowler, and F. Stern, Rev. Mod. Phys. **54**, 437 (1982)
[20] *The quantum Hall effect*, edited by R. Prange and S. Girvin (Springer-Verlag, Berlin, 1990).
[21] V. Gudmundson and R. Gerhards, Phys. Rev. B **35**, 8005 (1987).
[22] E. Stahl, D. Weiss, G. Weimann, K. von Klitzing, and K. Ploog, J. Phys. C **18**, L783 (1985).

[23] J. Kinoshita, K. Inagaki, C. Yamanouchi, K. Yoshihiro, J. Wakabayashi, and S. Kawaji, Proc. 2nd Intern. Symp. on Found. of Quantum Mechanics in the light of new technology, Sept. 1986, Phys. Soc. of Japan (1987).
[24] K. Kuchar, R. Meisels, G. Weimann, and W. Schlapp, Phys. Rev. B **33**, 2965 (1986).
[25] R.R. Gerhardts, R.J. Haug, and K. Ploog, J. de Phys. Colloques C **5**, 227 (1987).
[26] L. Landau, Z. Phys. **64**, 629 (1930).
[27] J. Zak, Phys. Rev. A **134**, 1602 (1964).
[28] P.W. Anderson, Phys. Rev. **109**, 1492 (1958).
[29] E. Abrahams, P.W. Anderson, D. Licciardello, and T. Ramakrishnan, Phys. Rev. Lett. **42**, 673 (1979).
[30] I. Goldsheid, S. Molcanov, and L. Pastur, Funct. Anal. Appl. **11**, 1 (1977).
[31] H. Kunz and B. Souillard, Commun. Math. Phys. **78**, 201 (1980).
[32] R. Carmona and J. Lacroix, *Spectral theory of random Schrödinger operators* (Birkhäuser, Boston, 1990).
[33] L. Pastur and A. Figotin, *Spectra of Random and Almost-Periodic Operators* (Springer-Verlag, Berlin, 1991).
[34] H. Cycon, R. Froese, W. Kirsch, and B. Simon, *Topics in the Theory of Schrödinger Operators* (Springer-Verlag, Berlin, 1987).
[35] J. Fröhlich and T. Spencer, Commun. Math. Phys. **88**, 151 (1983).
[36] F. Martinelli, E. Scoppola, J. Fröhlich, and T. Spencer, Commun. Math. Phys. **101**, 21 (1985).
[37] F. Delyon, Y. Levy, and B. Souillard, Commun. Math. Phys. **100**, 463 (1985).
[38] B. Simon and T. Wolff, Commun. Pure Appl. Math. **39**, 75 (1986).
[39] J. Bellissard, D. Grempel, F. Martinelli, and E. Scoppola, Phys. Rev. B **33**, 641 (1986).
[40] M. Aizenman and S. Molchanov, Commun. Math. Phys. **157**, 245 (1993).
[41] D.J. Thouless, Phys. Rep. C **13**, 93 (1974).
[42] S. Hikami, Nucl. Phys. B **215**, 555 (1983).
[43] J. Pichard and G. Sarma, J. Phys. C **14**, L127 and L617 (1981).
[44] A. MacKinnon and B. Kramer, Phys. Rev. Lett. **47**, 1546 (1981).
[45] A. MacKinnon and B. Kramer, Z. Phys. B **53**, 1 (1983).
[46] B. Huckestein and B. Kramer, Phys. Rev. Lett. **64**, 1437 (1990).
[47] W. Kirsch and F. Martinelli, J. Phys. A **15**, 2139 (1982).
[48] L. Schweitzer, B. Kramer, and A. MacKinnon, J. Phys. C **17**, 4111 (1984).
[49] R. Rammal, G. Toulouse, M. Jaekel, and B. Halperin, Phys. Rev. B **27**, 5142 (1983).
[50] Y. Hatsugai, Phys. Rev. B **48**, 2185 (1993).
[51] B.I. Halperin, Phys. Rev. B **25**, 2185 (1982).
[52] P. Harper, Proc. Phys. Soc. London, Ser. A **265**, 317 (1955).
[53] J. Bellissard and B. Simon, J. Funct. Anal. **49**, 191 (1982).
[54] J. Avron, P.H.M. v. Mouche, and B. Simon, Commun. Math. Phys. **132**, 103 (1990).
[55] M. Choi, G. Elliott, and N. Yui, Invent. Math. **99**, 225 (1990).
[56] R. Kubo, *Statistical mechanics* (North Holland, Amsterdam, 1965).
[57] H. Lawson and M. Michelson, *Spin Geometry* (Princeton University, Princeton, 1989).
[58] F. Hirzebruch, *Topological Methods in algebraic geometry* (Springer-Verlag, Berlin, 1966).
[59] J. Dixmier, Comptes rendus de l'académie des Sciences, 1107 (1966).
[60] A. Connes, *Commun. Math. Phys.* **117**, 673 (1988).
[61] J. Bellissard, in *From number theory to physics*, edited by M. Waldschmidt, P. Moussa, J. Luck, and C. Itzykson (Springer-Verlag, Berlin, 1991).
[62] M. Atiyah, *K-Theory* (Benjamin, New York, 1967).
[63] M.A. Shubin, Russ. Math. Surveys **34**, 109 (1979).
[64] Work in progress.
[65] A. Connes, in *Lecture Notes in Math.*, Vol. 725 (Springer-Verlag, Berlin, 1978).
[66] S. Sakai, *C*-Algebras and W*-Algebras* (Springer-Verlag, Berlin, 1971).
[67] B. Simon, *The $P(\phi)_2$ Euclidian (Quantum) Field theory* (Princeton University, Princeton, 1974).
[68] D. Pines and P. Nozières, *The Theory of Quantum Liquids*, Vol. 1 (Addison-Wesley, Reading, MA, 1989).
[69] T. Kato, *Perturbation theory for linear operators* (Springer-Verlag, Berlin, 1966).
[70] G. Casati, B.V. Chirikov, F.M. Izrailev, and J. Ford, in *Stochastic Behaviour in Classical and Quantum Hamiltonian Systems*, *Lecture Notes in Physics*, Vol. 93 (Springer-Verlag, Berlin, 1979).
[71] I. Guarneri, Lett. Nuovo Cimento **40**, 171 (1984).
[72] T. Takamasu, S. Komiyama, S. Hiyamizu, and S. Sasa, Surf. Sci. **170**, 202 (1986).
[73] D. Polyakov and B. Shklovskii, Phys. Rev. B **48**, 11167 (1993).
[74] A. Efros and B. Shklovskii, *Elctronic properties of doped semiconductors* (Springer, Berlin, 1984).
[75] G. Ebert, K. von Klitzing, C. Probst, E. Schuberth, K. Ploog, and G. Weimann, Solid State Commun. **45**, 625 (1983).
[76] B. Simon, *Trace ideals and their applications* (Cambridge University, Cambridge, 1979).
[77] G. Pedersen, *C*-algebras and their automorphism groups* (Academic, London, 1979).
[78] B. Blackadar, *K-Theory for Operator Algebras* (Springer-Verlag, Berlin, 1986).
[79] B. Fedosov, Funct. Anal. App. **4**, 339 (1967).
[80] J. Bellissard, *Noncommutative methods in semiclassical analysis*, to appear in *Lecture Notes in Math.*
[81] L. Pastur, Limits of infinite interaction radius, dimensionality and the number of components for randoms operators with off-diagonal randomness. Preprint #5, 1992/93, Institut Mittag-Leffler (1993).
[82] L. Pastur, Commun. Math. Phys. **75**, 179 (1980).

[83] I. Guarneri, private communication (1987).
[84] M. Aizenman, "Localization at Weak Disorder: Some Elementary Bounds," Princeton preprint (1993).
[85] S. Kotani, Cont. Math **50**, 277 (1986).
[86] S. Nakamura and J. Bellissard, Commun. Math. Phys. **131**, 283 (1990).
[87] Z. Tešanović, F. Axel, and B. Halperin, Phys. Rev. B **39**, 8525 (1989).
[88] S.A. Trugman, Phys. Rev. B **27**, 7539 (1983).
[89] G. Mil'nikov and I. Sokolov, P. Z. Eksp. Teor. Fiz. **48**, 494 (1988).
[90] J. Edwards and D.J. Thouless, J. Phys. C **5**, 807 (1972).
[91] D.P. Avoras, R.N. Bhatt, F.D.M. Haldane, P.B. Littlewood, and R. Rammal, Phys. Rev. Lett. **60**, 619 (1988).
[92] T. Ando, Surf. Sci. **196**, 120 (1988).
[93] Y. Huo and R. Bhatt, Phys. Rev. Lett. **68**, 1375 (1992).
[94] D. Liu and S. Das Sarma, Phys. Rev. B **49**, 2677 (1994).
[95] J.T. Chalker and P.D. Coddington, J. Phys. C **21**, 2665 (1988).
[96] H. Aoki, Surf. Sci. **263**, 137 (1992).
[97] Y. Tan, "Localization and quantum Hall effect in a two-dimensional periodic potential," Washington preprint (July 1993).
[98] H. Aoki and T. Ando, Phys. Rev. Lett **54**, 831 (1985).
[99] D. Hofstadter, Phys. Rev. B **14**, 2239 (1976).
[100] P. Středa, J. Phys. C **15**, L717 (1982).
[101] D.C. Tsui, H.L. Störmer, and A.C. Gossard, Phys. Rev. Lett. **48**, 1559 (1982).
[102] F.D.M. Haldane, Phys. Rev. Lett. **51**, 605 (1983).
[103] B.I. Halperin, Phys. Rev. Lett. **52**, 1583 (1984).
[104] J.K. Jain, Phys. Rev. Lett. **63**, 199 (1989).
[105] O. Bratteli and D.W. Robinson, *Operator Algebras and Quantum Statistical Mechanics*, Vol. 2 (Springer-Verlag, Berlin, 1981).
[106] F.D.M. Haldane, in *The Quantum Hall Effect*, edited by R.E. Prange and S.M. Girvin (Springer-Verlag, Berlin, 1990).
[107] R.B. Laughlin, in *The Quantum Hall Effect*, edited by R.E. Prange and S.M. Girvin (Springer-Verlag, Berlin, 1990).
[108] X.G. Wen, Phys. Rev. B **43**, 11025 (1991).
[109] J. Fröhlich and T. Kerler, Nucl. Phys. B **354**, 369 (1991).
[110] J. Fröhlich and A. Zee, Nucl. Phys. B **364**, 517 (1991).
[111] J. Fröhlich, "Mathematical Aspects of the Quantum Hall Effect," ETH-TH/92-15 preprint (1992).
[112] J. Fröhlich and U.M. Studer, Commun. Math. Phys. **148**, 553 (1992).
[113] J. Fröhlich and U.M. Studer, "Incompressible Quantum fluids, Gauge-Invariance and Current Algebra," preprint (1992).
[114] J. Fröhlich and U.M. Studer, "Gauge Invariance and Current Algebra in Non-Relativistic Many-Body Theory," preprint (1992).
[115] P. Goddard and D. Olive, Int. J. Mod. Phys. A **1**, 303 (1986).
[116] D. Buchholz, G. Mack, and I. Todorov, in *The Algebraic Theory of Superselection Sectors: Introduction and Recent Results*, edited by D. Kastler (World Scientific, Singapore, 1990).
[117] S. H. Simon and B. I. Halperin, "Explanation for the Resistivity Law in Quantum Hall Systems," preprint Harvard, 1994.