

A Topological Look at the Quantum Hall Effect

The amazingly precise quantization of Hall conductance in a two-dimensional electron gas can be understood in terms of a topological invariant known as the Chern number.

Joseph E. Avron, Daniel Osadchy, and Ruedi Seiler

The story of the Hall effect begins with a mistake made by James Clerk Maxwell. In the first edition of his *Treatise on Electricity and Magnetism*, which appeared in 1873, Maxwell discussed the deflection of a current by a magnetic field. He then said: “It must be carefully remembered that the mechanical force which urges a conductor . . . acts, not on the electric current, but on the conductor which carries it.” If readers are puzzled by this assertion, they should be.

In 1878, Edwin Hall, a student at Johns Hopkins University, was reading Maxwell for a class taught by Henry Rowland. Hall asked Rowland about Maxwell’s remark. The professor replied that he “doubted the truth of Maxwell’s statement and had sometime before made a hasty experiment . . . though without success.”¹

Hall made a fresh start and designed a different experiment, aimed at measuring, instead, the magnetoresistance—that is, the change of the electrical resistance due to the magnetic field. As we now know, that is a much harder experiment, and it too failed. Maxwell appeared to be safe. Hall then decided to repeat Rowland’s experiment. Following his mentor’s suggestion, Hall replaced the original metal conducting bar with a thin gold leaf, to compensate for the weakness of the available magnetic field.

That did the trick. A schematic diagram of Hall’s setup for examining what is now known as the Hall effect is shown in figure 1. He found that—Maxwell to the contrary notwithstanding—the magnetic field permanently altered the charge distribution, thereby deflecting the galvanometer connected to the lateral edges of the conductor. The transverse potential difference between the edges is called the Hall voltage. The Hall conductance is essentially the longitudinal current divided by this transverse voltage.

The discovery earned Hall a position at Harvard. His paper came out in 1879, the year of Maxwell’s death at age 48. In the second edition of Maxwell’s book, which appeared in 1881, a polite footnote by the editor says: “Mr. Hall has discovered that a steady magnetic field does slightly alter the distribution of currents in most conductors so that the statement . . . must be regarded as only approximately true.”

Joseph Avron is a professor of physics at the Technion–Israel Institute of Technology, in Haifa. **Daniel Osadchy** is a former student of Avron’s at the Technion. **Ruedi Seiler** is a professor of mathematics at the Technical University of Berlin, in Germany.

It turned out that the magnitude, and even the sign, of the Hall voltage depends on the material properties of the conductor—the gold leaf in Hall’s experiment. That made the Hall effect an important diagnostic tool for investigating the carriers of electric current. Eventually it pointed to the

concept of positively charged holes as current carriers in solids. So Maxwell, even when he was wrong, inspired fruitful research.

The quantum Hall effect

A century later, the Hall effect was revived as a source of astonishing new physics. In 1980 at the Grenoble High Magnetic Field Laboratory in France, Klaus von Klitzing was studying the Hall conductance of a two-dimensional electron gas at very low temperatures. He discovered, in essence, that the Hall conductance, as a function of the strength of the magnetic field applied normal to the gas plane, exhibited a staircase sequence of wide plateaus.² With totally unanticipated precision, the successive values of the Hall conductance were integer multiples of a fundamental constant of nature,

$$e^2/h = 1/(25\,812.807\,572\,\Omega),$$

irrespective of the geometric details of the experiment or the imperfections of its materials (see figure 2). Von Klitzing was awarded the 1985 Nobel Prize in Physics for the discovery of this quantum Hall effect (see PHYSICS TODAY, December 1985, page 17), and its precision has provided metrologists a superior standard of electrical resistance.

The quantum Hall effect can also be regarded as an appealingly straightforward precision measurement of the fine-structure constant, $e^2/\hbar c$, yielding a value of $1/137.0360\,0300(270)$. The alternative of measuring the electron’s anomalous magnetic moment does give the fine-structure constant with somewhat greater precision. But that path has required a huge calculational effort—more than a thousand Feynman diagrams—frought with chances for small mistakes.³

How can we understand the remarkable precision of Hall quantization despite the imprecise characterization of the experimental materials? Different samples have different impurities, different geometries, and different electron concentrations. Among the theoretical developments spawned by this question has been the recognition that the Hall conductance at the plateaus has topological significance.^{4,5} It can be understood in terms of topological invariants known as Chern numbers. After some preliminaries, we explain what Chern numbers are and how they relate to the quantum Hall effect.

Laughlin’s argument

In a 1981 paper,⁶ theorist Robert Laughlin put forward an argument to explain von Klitzing’s discovery (see PHYSICS TODAY, June 1981, page 17). That argument played a major role in the development of the theory of what has

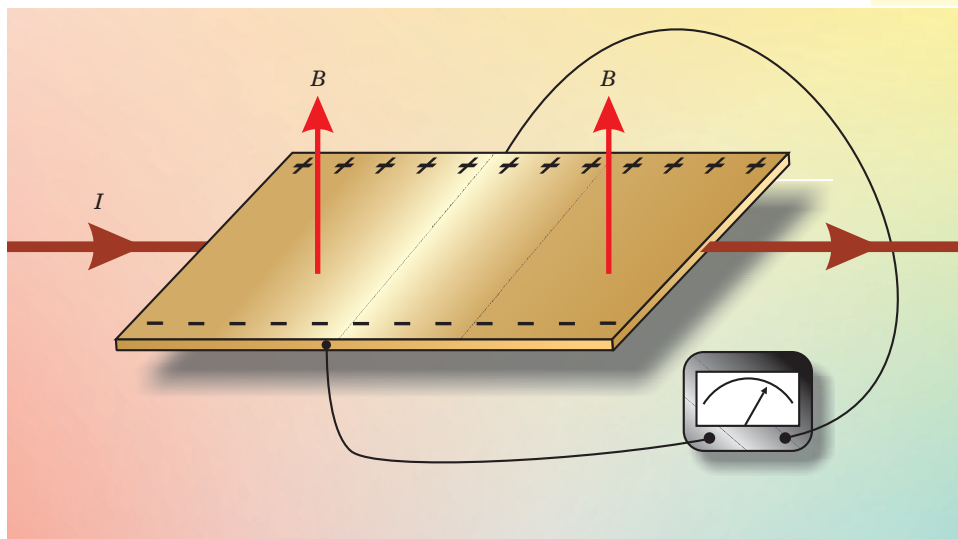


Figure 1. Edwin Hall's 1878 experiment was the first demonstration of the Hall effect. A magnetic field B normal to a gold leaf exerts a Lorentz force on a current I flowing longitudinally along the leaf. That force separates charges and builds up a transverse "Hall voltage" between the conductor's lateral edges. Hall detected this transverse voltage with a voltmeter that spanned the conductor's two edges.

come to be called the integer quantum Hall effect, and it deserves to be reexamined a quarter of a century later. (Laughlin also made seminal contributions to the understanding of the *fractional* quantum Hall effect, discovered in 1983. But that's another story.) His argument goes much of the way toward explaining the unexpected precision of the integral plateaus. But by our present understanding, it is short by one important step—namely, the inclusion of topological quantum numbers.

Laughlin considered a 2D electron gas cold enough so that quantum coherence holds throughout. It is then meaningful to speak of a wavefunction describing the system and its Hamiltonian evolution. Laughlin looked at the Hall effect as a quantum pump. He imagined the electron gas confined to a looped ribbon, as shown in figure 3, with a strong magnetic field normal to its surface. The two opposite edges of the ribbon are connected to separate electron reservoirs.

Laughlin then introduced a fictitious magnetic flux Φ threading the loop. The change in this flux drives the pump: Increasing the flux creates an electromotive force (emf) around the ring, which, by the classical Hall effect, results in the transfer of charge from one reservoir to the other. The Aharonov–Bohm principle tells us that the Hamiltonian describing the system is gauge invariant under flux changes by integral multiples of $\Phi_0 = hc/e$, the elementary quantum of magnetic flux (see PHYSICS TODAY,

January 1986, page 17). Therefore an adiabatic increase of Φ by a single flux quantum is a cycle of the pump.

An easy calculation shows that the charge transported between the reservoirs in one pump cycle, in units of the electron charge e , is the Hall conductance of the system in units of e^2/h , the quantum of Hall conductance. Therefore, if we can understand the precise quantization of the charge transported in one cycle of Laughlin's pump, we will understand the integer quantum Hall effect. In Laughlin's words, "By gauge invariance, adding Φ_0 maps the system back to itself, . . . [which results in] the transfer of n electrons." The quantization of Hall conductance is then implied.

We must ask, however, why the average transferred charge has to be an integral multiple of e , the charge of the electron. Classically, of course, an electron is either in reservoir A or B, but not in both. But why is that also true in a quantum mechanical system? Admittedly, even in quantum mechanics, a measurement of the number of electrons in a reservoir must be an integer, as must the transported charge. But in quantum mechanics, consecutive cycles of the pump may transport different amounts of charge.

Gauge invariance does require that, after a cycle, the pump is back in its original state. Doesn't that guarantee that the transported charge in different cycles must be the same? The answer is *no*. Only in classical mechanics does an exact reproduction of a prior state guarantee reproduction

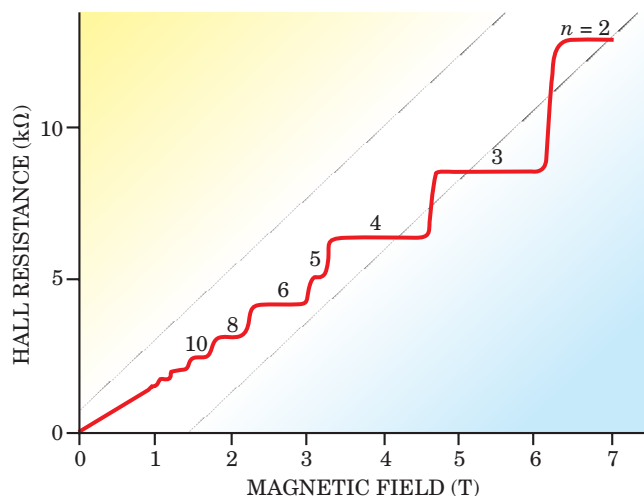
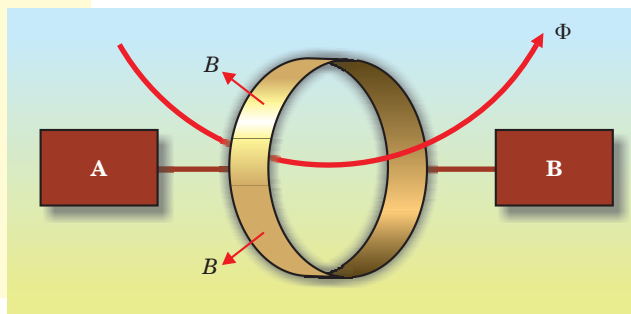


Figure 2. The integer quantum Hall effect. Plotting the Hall resistance (essentially the reciprocal of the Hall conductance) of a low-temperature two-dimensional electron gas against the strength of the imposed magnetic field normal to the gas plane, one finds a stairlike quantized sequence of Hall conductances very precisely equal to ne^2/h , where n is the integer that characterizes each plateau. The natural unit of resistance defined by this effect is about 26 k Ω . (Adapted from M. Paalanen, D. Tsui, A. Gossard, *Phys. Rev. B.* **25**, 5566 [1982].)

Figure 3. Robert Laughlin's 1981 gedanken experiment interprets the integer Quantum Hall effect as a quantum pump. Increasing the flux Φ that threads the conducting loop by a single flux quantum constitutes a cycle of the pump, transferring a quantized amount of charge between the two reservoirs, A and B, connected to the two edges of the conducting loop. The loop is everywhere subjected to a perpendicular magnetic field B .



of the prior measured result. In quantum mechanics, reproducing the state of the system does not necessarily reproduce the measurement outcome. So one cannot conclude from gauge invariance alone that the same number of electrons is transferred in every cycle of the pump.

Why, then, is the Hall conductance quantized? To complete the argument, one has to explain why the mean transferred charge, averaged over many pump cycles, is indeed quantized. That's where topological quantum numbers come into play: Chern numbers quantize averages.

Adiabatic curvature

In 1981, Michael Berry discovered that the phase accumulated by the wavefunction undergoing adiabatic evolution has a particular geometric component, now known as Berry's phase⁷ (see the article by Berry in *PHYSICS TODAY*, December 1990, page 34). To explain what Berry's phase is and its significance for the Hall effect, let's take a step back and review the notion of parallel transport in geometry.

In 1917, Tullio Levi-Civita developed the modern perspective on the geometry of surfaces based on Karl Friedrich Gauss's earlier work. In the Euclidean plane, there is an obvious notion of parallelism for vectors at different points. But that's not so on a curved surface, where there is no natural way to compare the directions of tangent vectors at different points. To compare directions, we need the notion of parallel transport.

For concreteness and simplicity, let us consider the surface of Earth—ignoring its rotation for the moment. The plane of a pendulum's swing defines a direction on the plane tangent to the surface. If the pendulum is moved slowly from one point to another, the propagation of that direction is a realization of parallel transport.

On the rotating planet, a Foucault pendulum is an example of parallel transport along a line of latitude. Parallel transport is an intriguing phenomenon, and the Foucault pendulum never fails to fascinate visitors to science museums. John Sullivan has created an interactive Web site that nicely illustrates parallel transport on a sphere.⁸ It shows how a vector can be transported parallel to itself and yet point in a different direction at the end of a round trip. That's what happens with the Foucault pendulum after 24 hours. Only at the poles and on the equator does the pendulum point in the same direction as it did 24 hours earlier.

The failure of parallel transport for closed paths is a hallmark of intrinsic curvature. In modern geometry, the local curvature of a surface is defined as the angular mismatch after the traversal of an infinitesimal closed loop, divided by the loop's area.

This notion of curvature extends to a wide range of other situations. In particular, it lets us introduce curvature into quantum mechanics. Consider a quantum Hamiltonian $H(\Phi, \theta)$ that depends on two angular parameters. The parameters play a role analogous to the spherical coordinates on Earth's surface. Suppose that the Hamiltonian has a nondegenerate ground state at energy zero. Let

$e^{i\alpha} |\psi(\Phi, \theta)\rangle$ denote the ground state. We are free to choose α as we please; it is the analog of the pendulum's initial direction.

Consider now a closed loop in the parameter space. If the parameters are varied slowly, we can use the time-dependent Schrödinger equation to transport the ground state. The failure of parallel transport around a closed loop is measured by Berry's phase. In this case, the local adiabatic curvature K of the bundle of ground states in the parameter space, defined as the limit of the Berry phase mismatch divided by the loop area, turns out to be

$$K = 2 \operatorname{Im} \langle \partial_{\Phi} \psi | \partial_{\theta} \psi \rangle. \quad (1)$$

Hall conductance as curvature

The Hall conductance can be thought of as a curvature. To see why, we identify the two angular parameters on which the Hall-effect Hamiltonian depends. One of them, Φ , is associated with the emf that drives the Hall current in figure 3. The second parameter, θ , is related to the ammeter that measures the Hall current. More precisely, θ is chosen in such a way that the Hall current takes the form: $I = c \partial_{\theta} H(\Phi, \theta)$. One can treat both Φ and θ as angular parameters because, by gauge invariance, the Hamiltonian is periodic in both, with period Φ_0 .

If Φ varies slowly and the ground-state energy is independent of Φ (and strictly below that of the first excited state), the Schrödinger equation gives

$$\langle \psi | I | \psi \rangle = \hbar c K \dot{\Phi} \quad (2)$$

for the expectation value of the Hall current, where K is the adiabatic curvature given by equation 1.

Equation 2 gives a linear relation between the expectation value of the Hall current and the driving emf, Φ/c , generated by the time-varying flux tube that threads the loop. The Hall conductance is therefore $\hbar c^2 K$. That relation establishes the geometric interpretation of the Hall conductance as curvature.

Ludwig Boltzmann is reputed to have said that elegance is for tailors. The geometric interpretation of the Hall conductance as curvature is clearly elegant. But is there more to it than elegance? There is: Geometry links the Hall conductance with topological invariants. Topology, therefore, is our next topic.

Chern numbers

Geometry and topology are intimately related. Let us recall this relation in the familiar setting of surfaces. A remarkable relation between geometry and topology is the formula by Gauss and Charles Bonnet:

$$\frac{1}{2\pi} \int_S K \, dA = 2(1 - g). \quad (3)$$

The integral is over a surface S without a boundary, like the torus in figure 4, and K is the local curvature of the surface. Therefore, $K \, dA$ is the angular mismatch of par-

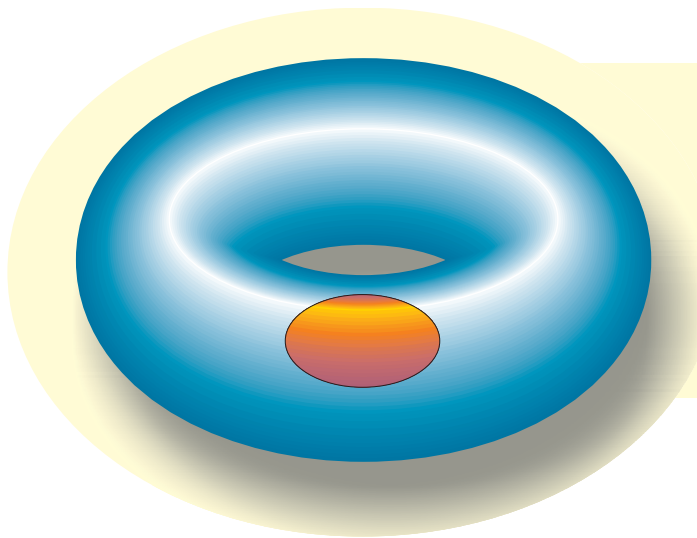


Figure 4. The Gauss–Bonnet formula (equation 3) is illustrated here by a toroidal surface with one handle. The local curvature K is positive on those portions of the surface that resemble a sphere and negative on those, near the hole, that resemble a saddle. Because g , the number of handles, equals one, the integral of the curvature over the entire surface vanishes. One can make Shiing-shen Chern’s quantum generalization of the Gauss–Bonnet formula plausible by considering the angular mismatch of parallel transport after a circuit around the small red patch in the figure.

allel transport around the small loop enclosing the area dA . The left side of the Gauss–Bonnet equation is geometric and not quantized a priori. But the right side is manifestly quantized; the integer g is the number of handles characterizing the topology of S . (For the torus in figure 4, $g = 1$.) So, if we change K arbitrarily by denting the surface—so long as we do not punch through new handles—the quantized right side of equation 3 does not change.

The Gauss–Bonnet formula has a well-known modern generalization, due to Shiing-shen Chern. The Gauss–Bonnet–Chern formula applies not only to the geometry of surfaces but also to the geometry of the eigenstates parameterized by Φ and θ . It looks exactly like equation 3, except that K is now the adiabatic curvature of equation 1, and the surface S is a torus parameterized by the two fluxes Φ and θ . The right-hand side of the Chern’s generalization of equation 3 is still an integer. Indeed, it is the so-called Chern number. But there is a difference: It is not necessarily an even integer, and g can no longer be interpreted as a tally of handles.

But if the Chern number doesn’t count handles, why does it have to be an integer? One can see why by considering the failure of parallel transport around the little loop in figure 4. The loop can be thought of as the boundary of the small red “inside” area, but also as the boundary of the “outside” area that makes up the rest of the toroidal surface. The phase mismatch of parallel transport can then be calculated as an integral of the curvature either over the inside area or the outside area. And the two integrals must agree up to an integral multiple of 2π . The difference between them, divided by 2π , is the Chern number. As one shrinks the small red area to zero, one readily sees that the Chern number is given by the integral in equation 3.

The Chern number is topological in the sense that it is invariant under small deformations of the Hamiltonian. Small changes of the Hamiltonian result in small changes of the adiabatic curvature and therefore, one might think, in a small change in the Chern number. But because the Chern number is an integer, it can’t change at all if it has to change continuously. We conclude, therefore, that a graph of the Chern number must have plateaus.

But then, how does the Chern number change from one plateau to the next? Large deformations of the Hamiltonian can cause the ground state to cross over other eigenstates. When such a “level crossing” happens in a quantum Hall system, the adiabatic curvature K diverges and the Chern number is no longer well defined. Transitions between Chern-number plateaus take place at level crossings.

Returning to Laughlin’s argument, we now see that, in fundamental units, the Hall conductance and therefore the average charge transport in Laughlin’s gedanken experiment are Chern numbers. That explains why the transported charge, averaged over many pump cycles, is quantized. And thus it explains the surprisingly precise plateau structure discovered in 1980 by von Klitzing.

The Hofstadter model

The glory of Chern numbers in the Hall effect emerges from a theoretical model investigated in 1976 by Douglas Hofstadter,⁹ three years before the publication of his widely read book *Gödel, Escher, Bach: An Eternal Golden Braid*. The model, which was first analyzed in the 1950s by Rudolph Peierls and his student P. G. Harper, describes independent electrons on a 2D lattice, acted on by a homogeneous magnetic field. The model is interesting for several reasons. First, its rich conjectural scheme was experimentally realized in 2001 by Christian Albrecht, von Klitzing, and coworkers in a 2D electron gas in a superlattice potential.¹⁰ The experiment verified the remarkably detailed Hofstadter-model predictions of David Thouless and coworkers.⁴ Second, the only known way of computing the Hall conductance in Hofstadter’s model is with Chern numbers. Third, the model provides a dazzling structure built of Chern numbers.

The thermodynamic properties of the 2D electron gas in Hofstadter’s model are determined by three parameters: the magnetic flux, the temperature, and the chemical potential, which fixes the electron density. The most interesting case, shown in figure 5, occurs at zero temperature.

The figure, sometimes called Hofstadter’s butterfly, is the model’s phase diagram. It represents all the thermodynamic phases of the 2D electron gas that emerge as one varies the chemical potential and the magnetic field. Unlike the phase diagrams of simple thermodynamic systems that have only a few phases—say liquid, solid, and gas—the Hofstadter model has infinitely many different phases. Each phase is characterized by its integral Hall conductance, and all integer values are allowed. These different Chern integers are represented in the figure by different colors. The figure was made using the equations of Thouless and company.⁴

How is it that something as simple as the Hofstadter model⁹ turns out to generate such a magnificently complex structure? The cause is, as is so often the case in physics, the phenomenon of “frustration.” Frustration occurs here because two different area scales are in competition. One characteristic area is the unit cell of the superlattice in Albrecht’s experiment. The other area scale is Φ_0/B , associated with the unit of quantum flux.

The character of the problem changes according to whether or not the two areas are commensurable. We say they are incommensurable when their ratio is an irrational

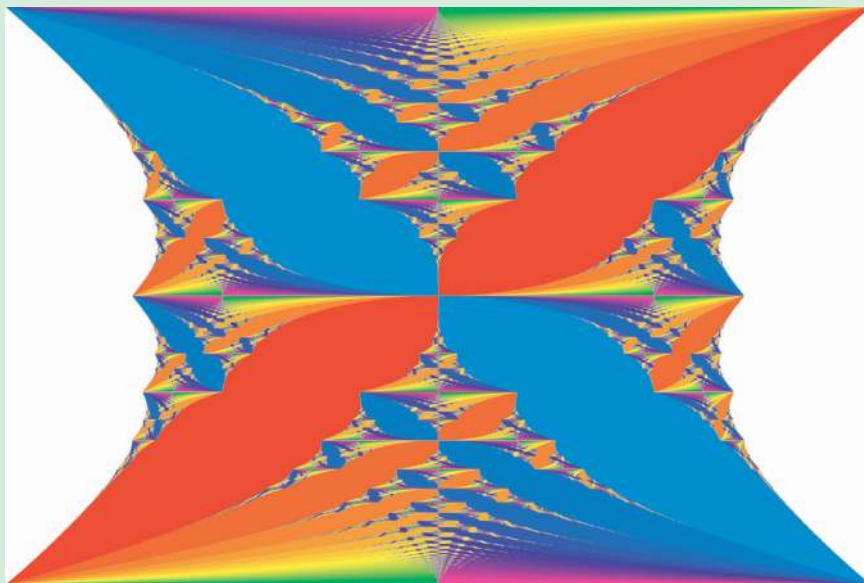


Figure 5. The Hofstadter “butterfly” is a fractal phase diagram of Douglas Hofstadter’s model of a quantum Hall system in a periodic potential. Each of the infinite number of phases is characterized by the Chern number of its Hall conductance and displayed with a different color. The horizontal axis gives the chemical potential, which determines the electron density. The left side corresponds to zero electron density. The vertical axis, indicating magnetic flux through the system, goes from zero at the bottom to one flux quantum per unit cell of the crystal lattice at the top. At higher fluxes, the phase diagram repeats periodically. Warm colors indicate positive Chern numbers; cool colors, negative numbers. Only on the extreme left and right do we see white spaces, which indicate Chern number zero. (Adapted from D. Osadchy, J. Avron, *J. Math. Phys.* **42**, 5665 [2001].)

number. Hofstadter’s 1976 paper begins by asking how the irrationality of a parameter could possibly affect an experimental outcome. Any such effect would seem to require infinite experimental resolution. The moral of the Hofstadter model is that commensurability has stable, experimentally realizable consequences in fractal-like structures. Rough experiments can resolve gross features like the large butterfly wings. Increasingly finer experiments eventually resolve finer and finer structures—the tiny wings—at the smallest energy scales.

Beyond elegance

Elegant results sometimes come at a price, and that is the case here: For the Chern-number theory to apply to the quantum Hall effect, a number of restrictive physical conditions must hold. We have mentioned only a few of them here. (Much of what we know thus far about the theory can be traced from references 11 and 12.) And we have omitted an interesting, alternative topological interpretation of the Hall conductance by an analog of the Chern number that arises in the context of a new branch of mathematics known as noncommutative geometry.^{5,13}

The geometric and topological ideas developed in the context of the quantum Hall effect have had an impact on an emerging direction in mesoscopic physics—quantum pumps.¹⁴ These pumps are quantum dots that can transport individual charges or spins between reservoirs.¹⁵ Here too, the transported charge can be related to an analog of the curvature.¹⁶

So far, we have barely mentioned the fractional quantum Hall effect. In 1983, Horst Stormer, Daniel Tsui, and coworkers at AT&T Bell Laboratories discovered that, in 2D electron gases confined at very high-quality semiconductor interfaces, the Hall conductance develops plateaus at simple *fractional* values of the Hall conductance. That unexpected discovery required a revision of the quantum Hall paradigm to a new theory in which electron–electron interaction plays a central role. For his theory of the fractional quantum Hall effect,¹¹ Laughlin shared the 1998 Nobel prize with Stormer and Tsui (see PHYSICS TODAY, December 1998, page 17.)

Laughlin’s 1983 theory of the fractional effect has evolved into a large body of knowledge.¹⁷ Chern numbers do not play a role in that theory. But topology does (see the article by J. K. Jain in PHYSICS TODAY, April 2000, page

39). A topological field theory based on the work of Chern and James Simon plays a significant role in recent formulations of the theory of the fractional quantum Hall effect.

We thank Fred Jegerlehner, Steve Lipson, Uri Sivan, and David Thouless for helpful comments. This work is supported by the Technion fund for the promotion of research and by grants from the European Union, the German Research Community, and the Swiss National Science Foundation.

References

1. E. H. Hall, *Am. J. Math.* **2**, 287 (1879).
2. K. von Klitzing, G. Dorda, M. Pepper, *Phys. Rev. Lett.* **45**, 494 (1980).
3. T. Kinoshita, M. Noi, *Phys. Rev. Lett.* **90**, 021803 (2003).
4. D. Thouless, M. Kohmoto, M. Nightingale, M. den Nijs, *Phys. Rev. Lett.* **49**, 405, (1982).
5. J. Bellissard, A. van Elst, H. Schultz-Baldes, *J. Math. Phys.* **35**, 5373 (1994); J. Avron, R. Seiler, B. Simon, *Commun. Math. Phys.* **159**, 399 (1994).
6. R. Laughlin, *Phys. Rev. B* **23**, 5632 (1981).
7. M. V. Berry, *Proc. R. Soc. London Ser. A* **392**, 45 (1984).
8. See <http://torus.math.uiuc.edu/jms/java/dragsphere>.
9. D. R. Hofstadter, *Phys. Rev. B* **14**, 2239 (1976).
10. C. Albrecht et al., *Phys. Rev. Lett.* **86**, 147 (2001).
11. M. Stone, ed., *The Quantum Hall Effect*, World Scientific, Singapore (1992); D. J. Thouless, *Topological Quantum Numbers in Nonrelativistic Physics*, World Scientific, River Edge, N.J. (1998).
12. M. Aizenman, G. Graf, *J. Phys. A* **31**, 6783 (1998).
13. J. Madore, *An Introduction to Noncommutative Differential Geometry and Its Physical Applications*, 2nd ed., Cambridge U. Press, New York (1999).
14. D. Thouless, *Phys. Rev. B* **27**, 6083 (1983); Q. Niu, *Phys. Rev. Lett.* **64**, 1812 (1990).
15. M. Switkes, C. Marcus, K. Campman, A. Gossard, *Science* **283**, 1907 (1999).
16. P. Brouwer, *Phys. Rev. B* **58**, 10135 (1998); J. Avron, A. Elgart, G. Graf, L. Sadun, *Phys. Rev. Lett.* **87**, 236601 (2001).
17. S. Girvin, in *Topological Aspects of Low Dimensional Systems*, A. Comtet, T. Jolicœur, S. Ouvry, F. David, eds., EDP Sciences, Cambridge, Mass. (1999); J. Fröhlich, U. Studer, E. Thiran, in *Fluctuating Geometries in Statistical Mechanics and Field Theory*, F. David, P. Ginsparg, J. Zinn-Justin, eds., Elsevier, New York (1996). ■