Séminaire Poincaré 2 (2004) 53 – 74 **Séminaire Poincaré**

Introduction to the Fractional Quantum Hall Effect

Steven M. Girvin
Yale University
Sloane Physics Laboratory
New Haven, CT 06520 USA

1 Introduction

The quantum Hall effect (QHE) is one of the most remarkable condensed-matter phenomena discovered in the second half of the 20th century. It rivals superconductivity in its fundamental significance as a manifestation of quantum mechanics on macroscopic scales. The basic experimental observation for a two-dimensional electron gas subjected to a strong magnetic field is nearly vanishing dissipation

$$\sigma_{xx} \to 0 \tag{1}$$

and special values of the Hall conductance

$$\sigma_{xy} = \nu \frac{e^2}{h} \tag{2}$$

given by the quantum of electrical conductance (e^2/h) multiplied by a quantum number ν. This quantization is universal and independent of all microscopic details such as the type of semiconductor material, the purity of the sample, the precise value of the magnetic field, and so forth. As a result, the effect is now used to maintain (but not define) the standard of electrical resistance by metrology laboratories around the world. In addition, since the speed of light is now defined, a measurement of e^2/h is equivalent to a measurement of the fine structure constant of fundamental importance in quantum electrodynamics.

Fig. (1) shows the remarkable transport data for a real device in the quantum Hall regime. Instead of a Hall resistivity which is simply a linear function of magnetic field, we see a series of so-called *Hall plateaus* in which ρ_{xy} is a universal constant

$$\rho_{xy} = -\frac{1}{\nu}\frac{h}{e^2} \tag{3}$$

independent of all microscopic details (including the precise value of the magnetic field). Associated with each of these plateaus is a dramatic decrease in the dissipative resistivity $\rho_{xx} \longrightarrow 0$ which drops as much as 13 orders of magnitude in the plateau regions. Clearly the system is undergoing some sort of sequence of phase transitions into highly idealized dissipationless states. Just as in a superconductor, the dissipationless state supports persistent currents.

In the so-called integer quantum Hall effect (IQHE) discovered by von Klitzing in 1980, the quantum number ν is a simple integer with a precision of about 10^{-10} and an absolute accuracy of about 10^{-8} (both being limited by our ability to do resistance metrology).

In 1982, Tsui, Störmer and Gossard discovered that in certain devices with reduced (but still non-zero) disorder, the quantum number ν could take on rational fractional values. This so-called fractional quantum Hall effect (FQHE) is the result of quite different underlying physics involving strong Coulomb interactions and correlations among the electrons. The particles condense into special quantum states whose excitations have the bizarre property of being described by fractional quantum numbers, including fractional charge and fractional statistics that are intermediate between ordinary Bose and Fermi statistics. The FQHE has proven to be a rich and surprising arena for the testing of our understanding of strongly correlated quantum systems. With a simple twist of a dial on her apparatus, the quantum Hall experimentalist can cause the electrons to condense

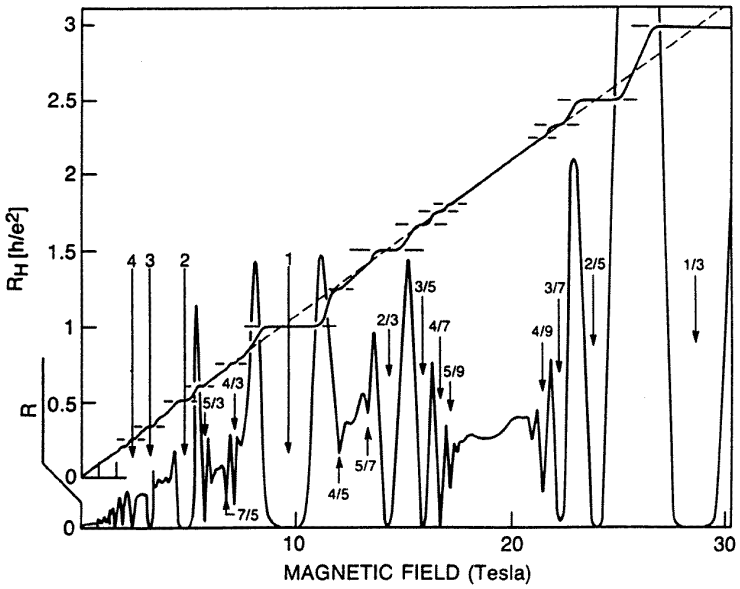

Figure 1: Integer and fractional quantum Hall transport data showing the plateau regions in the Hall resistance $R_{\rm H}$ and associated dips in the dissipative resistance R. The numbers indicate the Landau level filling factors at which various features occur. After ref. [1].

into a bewildering array of new 'vacua', each of which is described by a different quantum field theory. The novel order parameters describing each of these phases are completely unprecedented. A number of general reviews exist which the reader may be interested in consulting [2–10] The present lecture notes are based on the author's Les Houches Lectures. [11]

2 Fractional QHE

The free particle Hamiltonian an electron moving in a disorder-free two dimensional plane in a perpendicular magnetic field is

$$H = \frac{1}{2m}\Pi^2 \tag{4}$$

where

$$\vec{\Pi} \equiv \vec{p} + \frac{e}{c}\vec{A}(\vec{r}) \tag{5}$$

is the (mechanical) momentum. The magnetic field quenches the kinetic energy into discrete, massively degenerate Landau levels. In a sample of area $L_x L_y$, each Landau level has degeneracy equal to the number of flux quanta penetrating the sample

$$N_\Phi = L_x L_y \frac{B}{\Phi_0} = \frac{L_x L_y}{2\pi\ell^2} \tag{6}$$

where ℓ is the magnetic length defined by

$$\frac{1}{2\pi\ell^2} = \frac{B}{\Phi_0} \tag{7}$$

and $\Phi_0 = \frac{h}{e^2}$ is the quantum of flux. The quantum number ν in the quantized Hall coefficient turns out to be given by the Landau level filling factor

$$\nu = \frac{N}{N_\Phi}. \tag{8}$$

In the integer QHE the lowest ν Landau levels are completely occupied by electrons and the remainder at empty (at zero temperature). Under some circumstances of weak (but non-zero) disorder, quantized Hall plateaus appear which are characterized by simple rational fractional quantum numbers. For example, at magnetic fields three times larger than those at which the $\nu = 1$ integer filling factor plateau occurs, the lowest Landau level is only 1/3 occupied. The system ought to be below the percolation threshold (that is the electrons should be entirely localized by the weak random disorder potential) and hence be insulating. Instead a robust quantized Hall plateau is observed indicating that electrons can travel through the sample and that (since $\sigma_{xx} \longrightarrow 0$) there is an excitation gap (for all excitations except for the collective mode corresponding to uniform translation of the system which carries the current). This novel and quite unexpected physics is controlled by Coulomb repulsion between the electrons. It is best understood by first ignoring the disorder and trying to discover the nature of the special correlated many-body ground state into which the electrons condense when the filling factor is a rational fraction. Since the kinetic energy has been quenched, the Coulomb interaction has strong non-perturbative effects.

For reasons that will become clear later, it is convenient to analyze the problem in the so-called symmetric gauge

$$\vec{A} = -\frac{1}{2}\vec{r} \times \vec{B} \tag{9}$$

Unlike the Landau gauge which preserves translation symmetry in one direction, the symmetric gauge preserves rotational symmetry about the origin. Hence we anticipate that angular momentum (rather than one component of the linear momentum) will be a good quantum number in this gauge.

For simplicity we will restrict our attention to the lowest Landau level only and (simply to avoid some awkward minus signs) change the sign of the B field: $\vec{B} = -B\hat{z}$. With these restrictions, it is not hard to show that the solutions of the free-particle Schrödinger equation having definite angular momentum are

$$\varphi_m = \frac{1}{\sqrt{2\pi\ell^2 2^m m!}} z^m e^{-\frac{1}{4}|z|^2} \tag{10}$$

where $z = (x + iy)/\ell$ is a dimensionless complex number representing the position vector $\vec{r} \equiv (x, y)$ and $m \geq 0$ is an integer.

The angular momentum of these basis states is of course $\hbar m$. If we restrict our attention to the lowest Landau level, then there exists only one state with any given angular momentum and only non-negative values of m are allowed. This 'handedness' is a result of the chirality built into the problem by the magnetic field.

It seems rather peculiar that in the Landau gauge we have a continuous one-dimensional family of basis states corresponding to one component of conserved linear momentum for this two-dimensional problem. Now we find that in a different gauge, we have a discrete one dimensional label for the basis states! Nevertheless, we still end up with the correct density of states per unit area. To see this note that the peak value of $|\varphi_m|^2$ occurs at a radius of $R_{\text{peak}} = \sqrt{2m\ell^2}$. The area $2\pi\ell^2 m$ of a circle of this radius contains m flux quanta. Hence we obtain the standard result of one state per Landau level per quantum of flux penetrating the sample.

Because all the basis states are degenerate, any linear combination of them is also an allowed solution of the Schrödinger equation. Hence any function of the form [12]

$$\Psi(x, y) = f(z) e^{-\frac{1}{4}|z|^2} \tag{11}$$

is allowed so long as f is *analytic* in its argument. In particular, arbitrary polynomials of any degree N

$$f(z) = \prod_{j=1}^{N} (z - Z_j) \tag{12}$$

are allowed (at least in the thermodynamic limit) and are conveniently defined by the locations of their N zeros $\{Z_j; j = 1, 2, \ldots, N\}$. The fact that the Hilbert space for the lowest Landau level is the Hilbert space of analytic functions leads to some beautiful mathematics [11–13].

Another useful solution is the so-called coherent state which is a particular infinite order polynomial

$$f_\lambda(z) \equiv \frac{1}{\sqrt{2\pi\ell^2}} e^{\frac{1}{2}\lambda^* z} e^{-\frac{1}{4}\lambda^*\lambda}. \tag{13}$$

The wave function using this polynomial has the property that it is a narrow gaussian wave packet centered at the position defined by the complex number λ. Completing the square shows that the probability density is given by

$$|\Psi_\lambda|^2 = |f_\lambda|^2 e^{-\frac{1}{2}|z|^2} = \frac{1}{2\pi\ell^2} e^{-\frac{1}{2}|z-\lambda|^2} \tag{14}$$

This is the smallest wave packet that can be constructed from states within the lowest Landau level.

Because the kinetic energy is completely degenerate, the effect of Coulomb interactions among the particles is nontrivial. To develop a feel for the problem, let us begin by solving the two-body problem. Recall that the standard procedure is to take advantage of the rotational symmetry to write down a solution with the relative angular momentum of the particles being a good quantum number and then solve the Schrödinger equation for the radial part of the wave function. Here we find that the analyticity properties of the wave functions in the lowest Landau level greatly simplifies the situation. If we know the angular behavior of a wave function, analyticity uniquely defines the radial behavior. Thus for example for a single particle, knowing that the angular part of the wave function is $e^{im\theta}$, we know that the full wave function is guaranteed to uniquely be $r^m e^{im\theta} e^{-\frac{1}{4}|z|^2} = z^m e^{-\frac{1}{4}|z|^2}$.

Consider now the two body problem for particles with relative angular momentum m and center of mass angular momentum M. The *unique* analytic wave function is (ignoring normalization factors)

$$\Psi_{mM}(z_1, z_2) = (z_1 - z_2)^m (z_1 + z_2)^M e^{-\frac{1}{4}(|z_1|^2 + |z_2|^2)}. \tag{15}$$

If m and M are non-negative integers, then the prefactor of the exponential is simply a polynomial in the two arguments and so is a state made up of linear combinations of the degenerate one-body basis states φ_m given in eq. (10) and therefore lies in the lowest Landau level. Note that if the particles are spinless fermions then m must be odd to give the correct exchange symmetry. Remarkably, this is the exact (neglecting Landau level mixing) solution for the Schrödinger equation for *any* central potential $V(|z_1 - z_2|)$ acting between the two particles.[1] We do not need to solve any radial equation because of the powerful restrictions due to analyticity. There is only one state in the (lowest Landau level) Hilbert space with relative angular momentum m and center of mass angular momentum M. Hence (neglecting Landau level mixing) it is an exact eigenstate of *any* central potential. Ψ_{mM} is the exact answer independent of the Hamiltonian!

The corresponding energy eigenvalue v_m is independent of M and is referred to as the mth Haldane pseudopotential

$$v_m = \frac{\langle mM|V|mM\rangle}{\langle mM|mM\rangle}. \tag{16}$$

The Haldane pseudopotentials for the repulsive Coulomb potential are shown in Fig. (2). These discrete energy eigenstates represent bound states of the repulsive potential. If there were no magnetic field present, a repulsive potential would of course have only a continuous spectrum with no discrete bound states. However in the presence of the magnetic field, there are effectively bound states because the kinetic energy has been quenched. Ordinarily two particles that have a lot of potential energy because of their repulsive interaction can fly apart converting that potential energy into kinetic energy. Here however (neglecting Landau level mixing) the particles all have fixed kinetic energy. Hence particles that are repelling each other are stuck and can not escape from each other. One can view this semi-classically as the two particles orbiting each other under the influence of $\vec{E} \times \vec{B}$ drift with the Lorentz force preventing them from flying apart. In the

[1] Note that neglecting Landau level mixing is a poor approximation for strong potentials $V \gg \hbar\omega_c$ unless they are very smooth on the scale of the magnetic length.

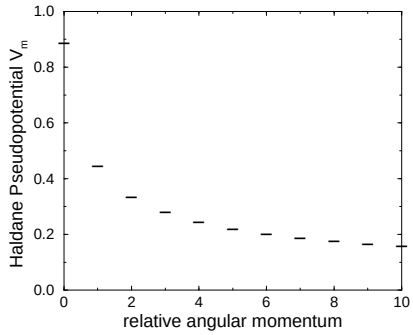

Figure 2: The Haldane pseudopotential V_m vs. relative angular momentum m for two particles interacting via the Coulomb interaction. Units are $e^2/\epsilon\ell$, where ϵ is the dielectric constant of the host semiconductor and the finite thickness of the quantum well has been neglected.

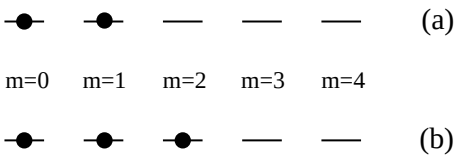

Figure 3: Orbital occupancies for the maximal density filled Landau level state with (a) two particles and (b) three particles. There are no particle labels here. In the Slater determinant wave function, the particles are labeled but a sum is taken over all possible permutations of the labels in order to antisymmetrize the wave function.

presence of an attractive potential the eigenvalues change sign, but of course the eigenfunctions remain exactly the same (since they are unique)!

The fact that a repulsive potential has a discrete spectrum for a pair of particles is (as we will shortly see) the central feature of the physics underlying the existence of an excitation gap in the fractional quantum Hall effect. One might hope that since we have found analyticity to uniquely determine the two-body eigenstates, we might be able to determine many-particle eigenstates exactly. The situation is complicated however by the fact that for three or more particles, the various relative angular momenta L_{12}, L_{13}, L_{23}, etc. do not all commute. Thus we can not write down general exact eigenstates. We will however be able to use the analyticity to great advantage and make exact statements for certain special cases.

2.1 The $\nu = 1$ many-body state

So far we have found the one- and two-body states. Our next task is to write down the wave function for a fully filled Landau level. We need to find

$$\psi[z] = f[z]\, e^{-\frac{1}{4}\sum_j |z_j|^2} \tag{17}$$

where $[z]$ stands for $(z_1, z_2, \ldots, z_N)$ and f is a polynomial representing the Slater determinant with all states occupied. Consider the simple example of two particles. We want one particle in the orbital φ_0 and one in φ_1, as illustrated schematically in Fig. (3a). Thus (again ignoring normalization)

$$\begin{aligned} f[z] &= \begin{vmatrix} (z_1)^0 & (z_2)^0 \\ (z_1)^1 & (z_2)^1 \end{vmatrix} = (z_1)^0(z_2)^1 - (z_2)^0(z_1)^1 \\ &= (z_2 - z_1) \end{aligned} \tag{18}$$

This is the lowest possible order polynomial that is antisymmetric. For the case of three particles we have (see Fig. (3b))

$$\begin{aligned} f[z] &= \begin{vmatrix} (z_1)^0 & (z_2)^0 & (z_3)^0 \\ (z_1)^1 & (z_2)^1 & (z_3)^1 \\ (z_1)^2 & (z_2)^2 & (z_3)^2 \end{vmatrix} = z_2 z_3^2 - z_3 z_2^2 - z_1^1 z_3^2 + z_3^1 z_1^2 + z_1 z_2^2 - z_2^1 z_1^2 \\ &= -(z_1 - z_2)(z_1 - z_3)(z_2 - z_3) \\ &= -\prod_{i<j}^{3}(z_i - z_j) \end{aligned} \tag{19}$$

This form for the Slater determinant is known as the Vandermonde polynomial. The overall minus sign is unimportant and we will drop it.

The single Slater determinant to fill the first N angular momentum states is a simple generalization of eq. (19)

$$f_N[z] = \prod_{i<j}^{N}(z_i - z_j). \tag{20}$$

To prove that this is true for general N, note that the polynomial is fully antisymmetric and the highest power of any z that appears is z^{N-1}. Thus the highest angular momentum state that is occupied is $m = N-1$. But since the antisymmetry guarantees that no two particles can be in the same state, all N states from $m = 0$ to $m = N-1$ must be occupied. This proves that we have the correct Slater determinant.

One can also use induction to show that the Vandermonde polynomial is the correct Slater determinant by writing

$$f_{N+1}(z) = f_N(z) \prod_{i=1}^{N}(z_i - z_{N+1}) \tag{21}$$

which can be shown to agree with the result of expanding the determinant of the $(N+1)\times(N+1)$ matrix in terms of the minors associated with the $(N+1)$st row or column.

Note that since the Vandermonde polynomial corresponds to the filled Landau level it is the unique state having the maximum density and hence is an exact eigenstate for any form of interaction among the particles (neglecting Landau level mixing and ignoring the degeneracy in the center of mass angular momentum).

The (unnormalized) probability distribution for particles in the filled Landau level state is

$$|\Psi[z]|^2 = \prod_{i<j}^{N}|z_i - z_j|^2 \; e^{-\frac{1}{2}\sum_{j=1}^{N}|z_j|^2}. \tag{22}$$

This seems like a rather complicated object about which it is hard to make any useful statements. It is clear that the polynomial term tries to keep the particles away from each other and gets larger as the particles spread out. It is also clear that the exponential term is small if the particles spread out too much. Such simple questions as, 'Is the density uniform?', seem hard to answer however.

It turns out that there is a beautiful analogy to plasma physics developed by R. B. Laughlin which sheds a great deal of light on the nature of this many particle probability distribution. To see how this works, let us pretend that the norm of the wave function

$$Z \equiv \int d^2 z_1 \ldots \int d^2 z_N \; |\psi_{[z]}|^2 \tag{23}$$

is the partition function of a classical statistical mechanics problem with Boltzmann weight

$$|\Psi[z]|^2 = e^{-\beta U_{\text{class}}} \tag{24}$$

where $\beta \equiv \frac{2}{m}$ and

$$U_{\text{class}} \equiv m^2 \sum_{i<j}(-\ln|z_i - z_j|) + \frac{m}{4}\sum_k |z_k|^2. \tag{25}$$

(The parameter $m = 1$ in the present case but we introduce it for later convenience.) It is perhaps not obvious at first glance that we have made tremendous progress, but we have. This is because $U_{\rm class}$ turns out to be the potential energy of a fake classical one-component plasma of particles of charge m in a uniform ('jellium') neutralizing background. Hence we can bring to bear well-developed intuition about classical plasma physics to study the properties of $|\Psi|^2$. Please remember however that all the statements we make here are about a particular wave function. There are no actual long-range logarithmic interactions in the quantum Hamiltonian for which this wave function is the approximate groundstate.

To understand this, let us first review the electrostatics of charges in three dimensions. For a charge Q particle in 3D, the surface integral of the electric field on a sphere of radius R surrounding the charge obeys

$$\int d\vec{A} \cdot \vec{E} = 4\pi Q. \tag{26}$$

Since the area of the sphere is $4\pi R^2$ we deduce

$$\vec{E}(\vec{r}) = Q\frac{\hat{r}}{r^2} \tag{27}$$

$$\varphi(\vec{r}) = \frac{Q}{r} \tag{28}$$

and

$$\vec{\nabla} \cdot \vec{E} = -\nabla^2 \varphi = 4\pi Q\, \delta^3(\vec{r}) \tag{29}$$

where φ is the electrostatic potential. Now consider a two-dimensional world where all the field lines are confined to a plane (or equivalently consider the electrostatics of infinitely long charged rods in 3D). The analogous equation for the line integral of the normal electric field on a *circle* of radius R is

$$\int d\vec{s} \cdot \vec{E} = 2\pi Q \tag{30}$$

where the 2π (instead of 4π) appears because the circumference of a circle is $2\pi R$ (and is analogous to $4\pi R^2$). Thus we find

$$\vec{E}(\vec{r}) = \frac{Q\hat{r}}{r} \tag{31}$$

$$\varphi(\vec{r}) = Q\left(-\ln\frac{r}{r_0}\right) \tag{32}$$

and the 2D version of Poisson's equation is

$$\vec{\nabla} \cdot \vec{E} = -\nabla^2 \varphi = 2\pi Q\, \delta^2(\vec{r}). \tag{33}$$

Here r_0 is an arbitrary scale factor whose value is immaterial since it only shifts φ by a constant.

We now see why the potential energy of interaction among a group of objects with charge m is

$$U_0 = m^2 \sum_{i<j} \left(-\ln|z_i - z_j|\right). \tag{34}$$

(Since $z = (x + iy)/\ell$ we are using $r_0 = \ell$.) This explains the first term in eq. (25).

To understand the second term notice that

$$-\nabla^2\, \frac{1}{4}|z|^2 = -\frac{1}{\ell^2} = 2\pi\rho_{\rm B} \tag{35}$$

where

$$\rho_{\rm B} \equiv -\frac{1}{2\pi\ell^2}. \tag{36}$$

Eq. (35) can be interpreted as Poisson's equation and tells us that $\frac{1}{4}|z|^2$ represents the electrostatic potential of a constant charge density $\rho_{\rm B}$. Thus the second term in eq. (25) is the energy of charge m objects interacting with this negative background.

Notice that $2\pi\ell^2$ is precisely the area containing one quantum of flux. Thus the background charge density is precisely B/Φ_0, the density of flux in units of the flux quantum.

The very long range forces in this fake plasma cost huge (fake) 'energy' unless the plasma is everywhere locally neutral (on length scales larger than the Debye screening length which in this case is comparable to the particle spacing). In order to be neutral, the density n of particles must obey

$$nm + \rho_{\rm B} = 0 \tag{37}$$

$$\Rightarrow \qquad n = \frac{1}{m}\,\frac{1}{2\pi\ell^2} \tag{38}$$

since each particle carries (fake) charge m. For our filled Landau level with $m = 1$, this is of course the correct answer for the density since every single-particle state is occupied and there is one state per quantum of flux.

We again emphasize that the energy of the fake plasma has nothing to do with the quantum Hamiltonian and the true energy. The plasma analogy is merely a statement about this particular choice of wave function. It says that the square of the wave function is very small (because $U_{\rm class}$ is large) for configurations in which the density deviates even a small amount from $1/(2\pi\ell^2)$. The electrons can in principle be found anywhere, but the overwhelming probability is that they are found in a configuration which is locally random (liquid-like) but with negligible density fluctuations on long length scales. We will discuss the nature of the typical configurations again further below in connection with Fig. (4).

When the fractional quantum Hall effect was discovered, Robert Laughlin realized that one could write down a many-body variational wave function at filling factor $\nu = 1/m$ by simply taking the mth power of the polynomial that describes the filled Landau level

$$f_N^m[z] = \prod_{i<j}^{N} (z_i - z_j)^m. \tag{39}$$

In order for this to remain analytic, m must be an integer. To preserve the antisymmetry m must be restricted to the odd integers. In the plasma analogy the particles now have fake charge m (rather than unity) and the density of electrons is $n = \frac{1}{m}\,\frac{1}{2\pi\ell^2}$ so the Landau level filling factor $\nu = \frac{1}{m} = \frac{1}{3}, \frac{1}{5}, \frac{1}{7}$, etc. (Later on, other wave functions were developed to describe more general states in the hierarchy of rational fractional filling factors at which quantized Hall plateaus were observed [2, 3, 5, 7, 8].)

The Laughlin wave function naturally builds in good correlations among the electrons because each particle sees an m-fold zero at the positions of all the other particles. The wave function vanishes extremely rapidly if any two particles approach each other, and this helps minimize the expectation value of the Coulomb energy.

Since the kinetic energy is fixed we need only concern ourselves with the expectation value of the potential energy for this variational wave function. Despite the fact that there are no adjustable variational parameters (other than m which controls the density) the Laughlin wave functions have proven to be very nearly exact for almost any realistic form of repulsive interaction. To understand how this can be so, it is instructive to consider a model for which this wave function actually is the exact ground state. Notice that the form of the wave function guarantees that every pair of particles has relative angular momentum greater than or equal to m. One should not make the mistake of thinking that every pair has relative angular momentum precisely equal to m. This would require the spatial separation between particles to be very nearly the same for every pair, which is of course impossible.

Suppose that we write the Hamiltonian in terms of the Haldane pseudopotentials

$$V = \sum_{m'=0}^{\infty} \sum_{i<j} v_{m'}\, P_{m'}(ij) \tag{40}$$

where $P_m(ij)$ is the projection operator which selects out states in which particles i and j have relative angular momentum m. If $P_{m'}(ij)$ and $P_{m''}(jk)$ commuted with each other things would be

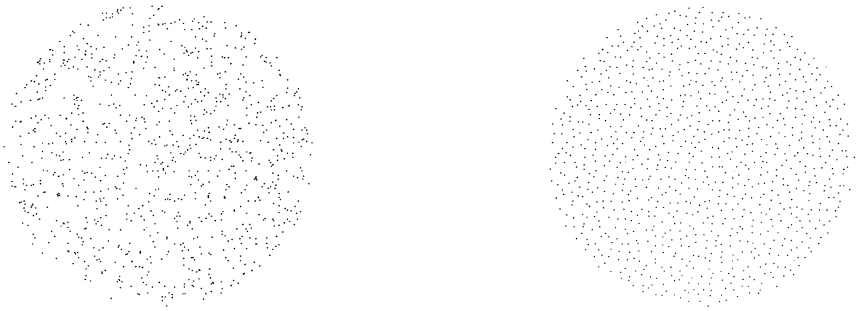

Figure 4: Comparison of typical configurations for a completely uncorrelated (Poisson) distribution of 1000 particles (left panel) to the distribution given by the Laughlin wave function for $m = 3$ (right panel). The latter is a snapshot taken during a Monte Carlo simulation of the distribution. The Monte Carlo procedure consists of proposing a random trial move of one of the particles to a new position. If this move increases the value of $|\Psi|^2$ it is always accepted. If the move decreases the value of $|\Psi|^2$ by a factor p, then the move is accepted with probability p. After equilibration of the plasma by a large number of such moves one finds that the configurations generated are distributed according to $|\Psi|^2$. (After R. B. Laughlin, Chap. 7 in [2].)

simple to solve, but this is not the case. However if we consider the case of a 'hard-core potential' defined by $v_{m'} = 0$ for $m' \geq m$, then clearly the mth Laughlin state is an exact, zero energy eigenstate

$$V\psi_m[z] = 0. \tag{41}$$

This follows from the fact that

$$P_{m'}(ij)\psi_m = 0 \tag{42}$$

for any $m' < m$ since every pair has relative angular momentum of at least m.

Because the relative angular momentum of a pair can change only in discrete (even integer) units, it turns out that this hard core model has an excitation gap. For example for $m = 3$, any excitation out of the Laughlin ground state necessarily weakens the nearly ideal correlations by forcing at least one pair of particles to have relative angular momentum 1 instead of 3 (or larger). This costs an excitation energy of order v_1.

This excitation gap is essential to the existence of dissipationless ($\sigma_{xx} = \rho_{xx} = 0$) current flow. In addition this gap means that the Laughlin state is stable against perturbations. Thus the difference between the Haldane pseudopotentials v_m for the Coulomb interaction and the pseudopotentials for the hard core model can be treated as a small perturbation (relative to the excitation gap). Numerical studies show that for realistic pseudopotentials the overlap between the true ground state and the Laughlin state is extremely good.

To get a better understanding of the correlations built into the Laughlin wave function it is useful to consider the snapshot in Fig. (4) which shows a typical configuration of particles in the Laughlin ground state (obtained from a Monte Carlo sampling of $|\psi|^2$) compared to a random (Poisson) distribution. Focussing first on the large scale features we see that density fluctuations at long wavelengths are severely suppressed in the Laughlin state. This is easily understood in terms of the plasma analogy and the desire for local neutrality. A simple estimate for the density fluctuations $\rho_{\vec{q}}$ at wave vector $\vec{q}$ can be obtained by noting that the fake plasma potential energy can be written (ignoring a constant associated with self-interactions being included)

$$U_{\text{class}} = \frac{1}{2L^2} \sum_{\vec{q}\neq 0} \frac{2\pi m^2}{q^2}\, \rho_{\vec{q}}\rho_{-\vec{q}} \tag{43}$$

where L^2 is the area of the system and $\frac{2\pi}{q^2}$ is the Fourier transform of the logarithmic potential (easily derived from $\nabla^2\left(-\ln(r)\right) = -2\pi\ \delta^2(\vec{r}\,)$). At long wavelengths ($q^2 \ll n$) it is legitimate

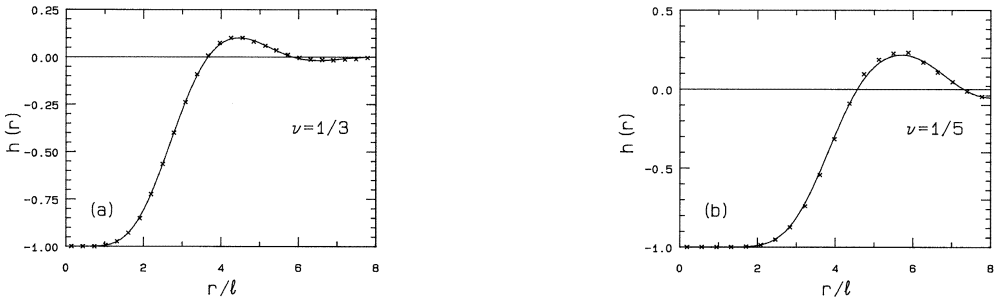

Figure 5: Plot of the two-point correlation function $h(r) \equiv 1 - g(r)$ for the Laughlin plasma with $\nu^{-1} = m = 3$ (left panel) and $m = 5$ (right panel). Notice that, unlike the result for $m = 1$ given in eq. (46), $g(r)$ exhibits the oscillatory behavior characteristic of a strongly coupled plasma with short-range solid-like local order.

to treat $\rho_{\vec{q}}$ as a collective coordinate of an elastic continuum. The distribution $e^{-\beta U_{\rm class}}$ of these coordinates is a gaussian and so obeys (taking into account the fact that $\rho_{-\vec{q}} = (\rho_{\vec{q}})^*$)

$$\langle \rho_{\vec{q}} \rho_{-\vec{q}} \rangle = L^2 \frac{q^2}{4\pi m}. \tag{44}$$

We clearly see that the long-range (fake) forces in the (fake) plasma strongly suppress long wavelength density fluctuations. We will return more to this point later when we study collective density wave excitations above the Laughlin ground state.

The density fluctuations on short length scales are best studied in real space. The radial correlation $g(r)$ function is a convenient object to consider. $g(r)$ tells us the density at r given that there is a particle at the origin

$$g(r) = \frac{N(N-1)}{n^2 Z} \int d^2 z_3 \ldots \int d^2 z_N \ |\psi(0, r, z_3, \ldots, z_N)|^2 \tag{45}$$

where $Z \equiv \langle \psi | \psi \rangle$, n is the density (assumed uniform) and the remaining factors account for all the different pairs of particles that could contribute. The factors of density are included in the denominator so that $\lim_{r\to\infty} g(r) = 1$.

Because the $m = 1$ state is a single Slater determinant $g(z)$ can be computed exactly

$$g(z) = 1 - e^{-\frac{1}{2}|z|^2}. \tag{46}$$

Fig. (5) shows numerical estimates of $h(r) \equiv 1 - g(r)$ for the cases $m = 3$ and 5. Notice that for the $\nu = 1/m$ state $g(z) \sim |z|^{2m}$ for small distances. Because of the strong suppression of density fluctuations at long wavelengths, $g(z)$ converges exponentially rapidly to unity at large distances. For $m > 1$, g develops oscillations indicative of solid-like correlations and, the plasma actually freezes[2] at $m \approx 65$. The Coulomb interaction energy can be expressed in terms of $g(z)$ as[3]

$$\frac{\langle \psi | V | \psi \rangle}{\langle \psi | \psi \rangle} = \frac{nN}{2} \int d^2 z \ \frac{e^2}{\epsilon |z|} \ [g(z) - 1] \tag{47}$$

where the (-1) term accounts for the neutralizing background and ϵ is the dielectric constant of the host semiconductor. We can interpret $g(z) - 1$ as the density of the 'exchange-correlation hole' surrounding each particle.

[2]That is, Monte Carlo simulation of $|\Psi|^2$ shows that the particles are most likely to be found in a crystalline configuration which breaks translation symmetry. Again we emphasize that this is a statement about the Laughlin variational wave function, not necessarily a statement about what the electrons actually do. It turns out that for $m \geq\sim 7$ the Laughlin wave function is no longer the best variational wave function. One can write down wave functions describing Wigner crystal states which have lower variational energy than the Laughlin liquid.

[3]This expression assumes a strictly zero thickness electron gas. Otherwise one must replace $\frac{e^2}{\epsilon|z|}$ by $\frac{e^2}{\epsilon} \int_{-\infty}^{+\infty} ds \frac{|F(s)|^2}{\sqrt{|z|^2+s^2}}$ where F is the wavefunction factor describing the quantum well bound state.

The correlation energies per particle for $m = 3$ and 5 are [14]

$$\frac{1}{N}\,\frac{\langle\psi_3|V|\psi_3\rangle}{\langle\psi_3|\psi_3\rangle} = -0.4100 \pm 0.0001 \tag{48}$$

and

$$\frac{1}{N}\,\frac{\langle\psi_5|V|\psi_5\rangle}{\langle\psi_5|\psi_5\rangle} = -0.3277 \pm 0.0002 \tag{49}$$

in units of $e^2/\epsilon\ell$ which is ≈ 161 K for $\epsilon = 12.8$ (the value in GaAs), $B = 10$T. For the filled Landau level ($m = 1$) the exchange energy is $-\sqrt{\frac{\pi}{8}}$ as can be seen from eqs. (46) and (47).

3 Neutral Collective Excitations

So far we have studied one particular variational wave function and found that it has good correlations built into it as graphically illustrated in Fig. 4. To further bolster the case that this wave function captures the physics of the fractional Hall effect we must now demonstrate that there is finite energy cost to produce excitations above this ground state. In this section we will study the neutral collective excitations. We will examine the charged excitations in the next section.

It turns out that the neutral excitations are phonon-like excitations similar to those in solids and in superfluid helium. We can therefore use a simple modification of Feynman's 'single mode approximation' (SMA) theory of the excitations in superfluid helium [15, 16].

By way of introduction let us start with the simple harmonic oscillator. The ground state is of the form

$$\psi_0(x) \sim e^{-\alpha x^2}. \tag{50}$$

Suppose we did not know the excited state and tried to make a variational ansatz for it. Normally we think of the variational method as applying only to ground states. However it is not hard to see that the first excited state energy is given by

$$\epsilon_1 = \min\left\{\frac{\langle\psi|H|\psi\rangle}{\langle\psi|\psi\rangle}\right\} \tag{51}$$

provided that we do the minimization over the set of states ψ which are constrained to be orthogonal to the ground state ψ_0. One simple way to produce a variational state which is automatically orthogonal to the ground state is to change the parity by multiplying by the first power of the coordinate

$$\psi_1(x) \sim x\, e^{-\alpha x^2}. \tag{52}$$

Variation with respect to α of course leads (in this special case) to the *exact* first excited state.

With this background let us now consider the case of phonons in superfluid ^{4}He. Feynman argued that because of the Bose statistics of the particles, there are no low-lying single-particle excitations. This is in stark contrast to a fermi gas which has a high density of low-lying excitations around the fermi surface. Feynman argued that the only low-lying excitations in ^{4}He are collective density oscillations that are well-described by the following family of variational wave functions (that has no adjustable parameters) labelled by the wave vector

$$\psi_{\vec{k}} = \frac{1}{\sqrt{N}}\,\rho_{\vec{k}}\,\Phi_0 \tag{53}$$

where Φ_0 is the exact ground state and

$$\rho_{\vec{k}} \equiv \sum_{j=1}^{N} e^{-i\vec{k}\cdot\vec{r}_j} \tag{54}$$

is the Fourier transform of the density. The physical picture behind this is that at long wavelengths the fluid acts like an elastic continuum and $\rho_{\vec{k}}$ can be treated as a generalized oscillator normal-mode coordinate. In this sense eq. (53) is then analogous to eq. (52). To see that $\psi_{\vec{k}}$ is orthogonal

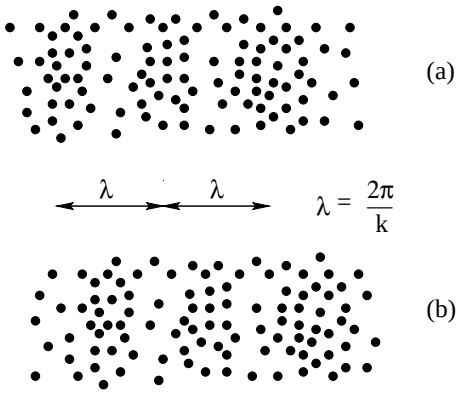

Figure 6: (a) Configuration of particles in which the Fourier transform of the density at wave vector k is non-zero. (b) The Fourier amplitude will have a similar magnitude for this configuration but a different phase.

to the ground state we simply note that

$$\begin{aligned} \langle\Phi_0|\psi_{\vec{k}}\rangle &= \frac{1}{\sqrt{N}}\,\langle\Phi_0|\rho_{\vec{k}}|\Phi_0\rangle \\ &= \frac{1}{\sqrt{N}}\int d^3R\; e^{-i\vec{k}\cdot\vec{R}}\;\langle\Phi_0|\rho(\vec{r})|\Phi_0\rangle. \end{aligned} \tag{55}$$

where

$$\rho(\vec{r}) \equiv \sum_{j=1}^{N} \delta^3(\vec{r}_j - \vec{R}) \tag{56}$$

is the density operator. If Φ_0 describes a translationally invariant liquid ground state then the Fourier transform of the mean density vanishes for $k \neq 0$.

There are several reasons why $\psi_{\vec{k}}$ is a good variational wave function, especially for small k. First, it contains the ground state as a factor. Hence it contains all the special correlations built into the ground state to make sure that the particles avoid close approaches to each other without paying a high price in kinetic energy. Second, $\psi_{\vec{k}}$ builds in the features we expect on physical grounds for a density wave. To see this, consider evaluating $\psi_{\vec{k}}$ for a configuration of the particles like that shown in Fig. (6a) which has a density modulation at wave vector $\vec{k}$. This is not a configuration that maximizes $|\Phi_0|^2$, but as long as the density modulation is not too large and the particles avoid close approaches, $|\Phi_0|^2$ will not fall too far below its maximum value. More importantly, $|\rho_{\vec{k}}|^2$ will be much larger than it would for a more nearly uniform distribution of positions. As a result $|\psi_{\vec{k}}|^2$ will be large and this will be a likely configuration of the particles in the excited state. For a configuration like that in Fig. (6b), the phase of $\rho_{\vec{k}}$ will shift but $|\psi_{\vec{k}}|^2$ will have the same magnitude. This is analogous to the parity change in the harmonic oscillator example. Because all different phases of the density wave are equally likely, $\rho_{\vec{k}}$ has a mean density which is uniform (translationally invariant).

This phonon mode should not be confused with the ordinary hydrodynamic sound mode in classical fluids. The latter occurs in a collision dominated regime $\omega\tau \ll 1$ in which collision-induced pressure provides the restoring force. The phonon mode described here by $\psi_{\vec{k}}$ is a low-lying eigenstate of the quantum Hamiltonian.

At larger wave vectors there is a so-called 'roton minimum' (see Fig. (7)) in the dispersion caused by the solid-like oscillations in the radial distribution function $g(r)$ similar to those shown in Fig. 5 for the Laughlin liquid. This minimum is in some crude sense a remnant of the zone boundary phonon of the crystal.

As we mentioned previously Feynman argued that in ^{4}He the Bose symmetry of the wave functions guarantees that unlike in Fermi systems, there is only a single low-lying mode, namely

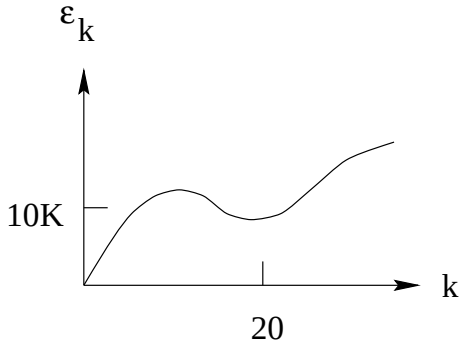

Figure 7: Schematic illustration of the phonon dispersion in superfluid liquid ^{4}He. For small wave vectors the dispersion is linear, as is expected for a gapless Goldstone mode. The roton minimum due to the peak in the static structure factor occurs at a wave vector k of approximately 20 in units of inverse Å. The roton energy is approximately 10 in units of Kelvins.

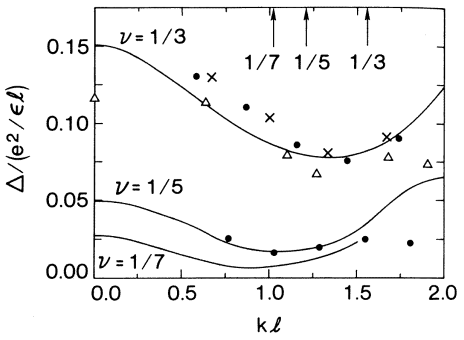

Figure 8: Comparison of the single mode approximation (SMA) prediction of the collective mode energy for filling factors $\nu = 1/3, 1/5, 1/7$ (solid lines) with small-system numerical results for N particles. Crosses indicate the $N = 7, \nu = 1/3$ spherical system, triangles indicate the $N = 6, \nu = 1/3$ hexagonal unit cell system results of Haldane and Rezayi [18]. Solid dots are for $N = 9, \nu = 1/3$ and $N = 7, \nu = 1/5$ spherical system calculations of Fano et al. [19] Arrows at the top indicate the magnitude of the reciprocal lattice vector of the Wigner crystal at the corresponding filling factor. Notice that unlike the phonon collective mode in superfluid helium shown in Fig. (7), the mode here is gapped.

the phonon density mode. The paucity of low-energy single particle excitations in boson systems is what helps make them superfluid–there are no dissipative channels for the current to decay into. Despite the fact that the quantum Hall system is made up of fermions, the behavior is also reminiscent of superfluidity since the current flow is dissipationless. Indeed, within the 'composite boson' picture, one views the FQHE ground state as a bose condensate [8, 9, 17].

It turns out that the SMA works extremely well in the FQHE as can be seen in Fig. (8). Because of the lack of density fluctuations at long wavelengths in the Laughlin ground state, the system is incompressible leading to a gap in the collective excitation spectrum at long wavelengths. This is quite different from the case of superfluid ^{4}He in which the mode is gapless. However like the case of the superfluid, this 'magnetophonon' mode has a 'magnetoroton' minimum at finite k as illustrated in Fig. (8). The figure also shows results from numerical exact diagonalization studies which demonstrate that the single mode approximation is extremely accurate. Note that the magnetoroton minimum occurs close to the position of the smallest reciprocal lattice vector in the Wigner crystal of the same density. In the crystal the phonon frequency would go exactly to zero at this point. (Recall that in a crystal the phonon dispersion curves have the periodicity of the reciprocal lattice.)

Because the oscillator strength is almost entirely in the cyclotron mode, the dipole matrix

element for coupling the collective excitations to light is very small. They have however been observed in Raman scattering [20] and found to have an energy gap in excellent quantitative agreement with the single mode approximation.

Finally we remark that these collective excitations are characterized by a well-defined wave vector $\vec{k}$ despite the presence of the strong magnetic field. This is only possible because they are charge neutral which allows one to define a gauge invariant conserved momentum [21].

4 Charged Excitations

Except for the fact that they are gapped, the neutral magnetophonon excitations are closely analogous to the phonon excitations in superfluid ^{4}He. We further pursue this analogy with a search for the analog of vortices in superfluid films. A vortex is a topological defect which is the quantum version of the familiar whirlpool. A reasonably good variational wave function for a vortex in a two-dimensional film of ^{4}He is

$$\psi^{\pm}_{\vec{R}} = \left\{ \prod_{j=1}^{N} f\left(|\vec{r}_j - \vec{R}|\right)\ e^{\pm i\theta(\vec{r}_j - \vec{R})} \right\} \Phi_0. \tag{57}$$

Here θ is the azimuthal angle that the particle's position makes relative to $\vec{R}$, the location of the vortex center. The function f vanishes as $\vec{r}$ approaches $\vec{R}$ and goes to unity far away. The choice of sign in the phase determines whether the vortex is right or left handed.

The interpretation of this wave function is the following. The vortex is a topological defect because if any particle is dragged around a closed loop surrounding $\vec{R}$, the phase of the wave function winds by $\pm 2\pi$. This phase gradient means that current is circulating around the core. Consider a large circle of radius ξ centered on $\vec{R}$. The phase change of 2π around the circle occurs in a distance $2\pi\xi$ so the local gradient seen by *every* particle is $\hat{\theta}/\xi$. We see that locally the center of mass momentum has been boosted by $\pm\frac{\hbar}{\xi}\ \hat{\theta}$ so that the current density of the whirlpool falls off inversely with distance from the core.[4] Near the core f falls to zero because of the 'centrifugal barrier' associated with this circulation. In a more accurate variational wave function the core would be treated slightly differently but the asymptotic large distance behavior would be unchanged.

What is the analog of all this for the lowest Landau level? For ψ^+ we see that every particle has its angular momentum boosted by one unit. In the lowest Landau level analyticity (in the symmetric gauge) requires us to replace $e^{i\theta}$ by $z = x + iy$. Thus we are led to the Laughlin 'quasi-hole' wave function

$$\psi^+_Z[z] = \prod_{j=1}^{N} (z_j - Z)\ \psi_m[z] \tag{58}$$

where Z is a complex number denoting the position of the vortex and ψ_m is the Laughlin wave function at filling factor $\nu = 1/m$. The corresponding antivortex ('quasi-electron' state) involves z_j^* suitably projected into the Hilbert space [11, 12]:

$$\psi^-_Z[z] = \prod_{j=1}^{N} \left(2\frac{\partial}{\partial z_j} - Z^*\right)\ \psi_m[z] \tag{59}$$

where as usual the derivatives act only on the polynomial part of ψ_m. All these derivatives make ψ^- somewhat difficult to work with. We will therefore concentrate on the quasi-hole state ψ^+. The origin of the names quasi-hole and quasi-electron will become clear shortly.

Unlike the case of a superfluid film, the presence of the vector potential allows these vortices to cost only a finite energy to produce and hence the electrical dissipation is always finite at any

[4] This slow algebraic decay of the current density means that the total kinetic energy of a single vortex diverges logarithmically with the size of the system. This in turn leads to the Kosterlitz Thouless phase transition in which pairs of vortices bind together below a critical temperature. As we will see below there is no corresponding finite temperature transition in a quantum Hall system.

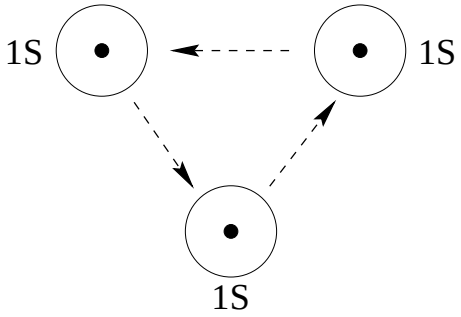

Figure 9: Illustration of an electron tunneling among the 1S orbitals of three protons. The tunneling is exponentially slow for large separations which leads to only exponentially small lifting of what would otherwise be a three-fold degenerate ground state.

non-zero temperature. There is no finite temperature transition into a superfluid state as in the Kosterlitz Thouless transition. From a field theoretic point of view, this is closely analogous to the Higg's mechanism [17].

Just as in our study of the Laughlin wave function, it is very useful to see how the plasma analogy works for the quasi-hole state

$$|\psi_Z^+|^2 = e^{-\beta U_{\rm class}}\ e^{-\beta V} \tag{60}$$

where $U_{\rm class}$ is given by eq. (25), $\beta = 2/m$ as before and

$$V \equiv m \sum_{j=1}^{N} \left(-\ln|z_j - Z|\right). \tag{61}$$

Thus we have the classical statistical mechanics of a one-component plasma of (fake) charge m objects seeing a neutralizing jellium background plus a new potential energy V representing the interaction of these objects with an 'impurity' located at Z and having unit charge.

Recall that the chief desire of the plasma is to maintain charge neutrality. Hence the plasma particles will be repelled from Z. Because the plasma particles have fake charge m, the screening cloud will have to have a net reduction of $1/m$ particles to screen the impurity. But this means that the quasi-hole has fractional fermion number! The (true) physical charge of the object is a fraction of the elementary charge

$$q^* = \frac{e}{m}. \tag{62}$$

This is very strange! How can we possibly have an elementary excitation carrying fractional charge in a system made up entirely of electrons? To understand this let us consider an example of another quantum system that seems to have fractional charge, but in reality doesn't. Imagine three protons arranged in an equilateral triangle as shown in Fig. (9). Let there be one electron in the system. In the spirit of the tight-binding model we consider only the 1S orbital on each of the three 'lattice sites'. The Bloch states are

$$\psi_k = \frac{1}{\sqrt{3}} \sum_{j=1}^{3} e^{ikj}\ |j\rangle \tag{63}$$

where $|j\rangle$ is the 1S orbital for the jth atom. The equilateral triangle is like a linear system of length 3 with periodic boundary conditions. Hence the allowed values of the wavevector are $\left\{k_\alpha = \frac{2\pi}{3}\alpha;\ \ \alpha = -1, 0, +1\right\}$. The energy eigenvalues are

$$\epsilon_{k_\alpha} = -E_{\rm 1S} - 2J\ \cos k_\alpha \tag{64}$$

where $E_{\rm 1S}$ is the isolated atom energy and $-J$ is the hopping matrix element related to the orbital overlap and is exponentially small for large separations of the atoms.

The projection operator that measures whether or not the particle is on site n is

$$P_n \equiv |n\rangle \langle n|. \tag{65}$$

Its expectation value in any of the three eigenstates is

$$\langle \psi_{k_\alpha} | P_n | \psi_{k_\alpha} \rangle = \frac{1}{3}. \tag{66}$$

This equation simply reflects the fact that as the particle tunnels from site to site it is equally likely to be found on any site. Hence it will, on average, be found on a particular site n only 1/3 of the time. The average electron number per site is thus 1/3. This however is a trivial example because the value of the measured charge is always an integer. Two-thirds of the time we measure zero and one third of the time we measure unity. This means that the charge *fluctuates*. One measure of the fluctuations is

$$\sqrt{\langle P_n^2 \rangle - \langle P_n \rangle^2} = \sqrt{\frac{1}{3} - \frac{1}{9}} = \frac{\sqrt{2}}{3}, \tag{67}$$

which shows that the fluctuations are larger than the mean value. This result is most easily obtained by noting $P_n^2 = P_n$.

A characteristic feature of this 'imposter' fractional charge $\frac{e}{m}$ that guarantees that it fluctuates is the existence in the spectrum of the Hamiltonian of a set of m nearly degenerate states. (In our toy example here, $m = 3$.) The characteristic time scale for the charge fluctuations is $\tau \sim \hbar/\Delta\epsilon$ where $\Delta\epsilon$ is the energy splitting of the quasi-degenerate manifold of states. In our tight-binding example $\tau \sim \hbar/J$ is the characteristic time it takes an electron to tunnel from the 1S orbital on one site to the next. As the separation between the sites increases this tunneling time grows exponentially large and the charge fluctuations become exponentially slow and thus easy to detect.

In a certain precise sense, the fractional charge of the Laughlin quasiparticles behaves very differently from this. An electron added at low energies to a $\nu = 1/3$ quantum Hall fluid breaks up into three charge 1/3 Laughlin quasiparticles. These quasiparticles can move arbitrarily far apart from each other[5] and yet no quasi-degenerate manifold of states appears. The excitation gap to the first excited state remains finite. The only degeneracy is that associated with the positions of the quasiparticles. If we imagine that there are three impurity potentials that pin down the positions of the three quasiparticles, then the state of the system is *uniquely* specified. Because there is no quasidegeneracy, we do not have to specify any more information other than the positions of the quasiparticles. Hence in a deep sense, they are true *elementary particles* whose fractional charge is a sharp quantum observable.

Of course, since the system is made up only of electrons, if we capture the charges in some region in a box, we will always get an integer number of electrons inside the box. However in order to close the box we have to locally destroy the Laughlin state. This will cost (at a minimum) the excitation gap. This may not seem important since the gap is small — only a few Kelvin or so. But imagine that the gap were an MeV or a GeV. Then we would have to build a particle accelerator to 'close the box' and probe the fluctuations in the charge. These fluctuations would be analogous to the ones seen in quantum electrodynamics at energies above $2m_e c^2$ where electron-positron pairs are produced during the measurement of charge form factors by means of a scattering experiment.

Put another way, the charge of the Laughlin quasiparticle fluctuates but only at high frequencies $\sim \Delta/\hbar$. If this frequency (which is $\sim$ 50GHz) is higher than the frequency response limit of our voltage probes, we will see no charge fluctuations. We can formalize this by writing a modified projection operator [22] for the charge on some site n by

$$P_n^{(\Omega)} \equiv P^\Omega \, P_n P^\Omega \tag{68}$$

where $P_n = |n\rangle \langle n|$ as before and

$$P^{(\Omega)} \equiv \theta(\Omega - H + E_0) \tag{69}$$

[5]Recall that unlike the case of vortices in superfluids, these objects are unconfined.

is the operator that projects onto the subset of eigenstates with excitation energies less than Ω. $P_n^{(\Omega)}$ thus represents a measurement with a high-frequency cutoff built in to represent the finite bandwidth of the detector. Returning to our tight-binding example, consider the situation where J is large enough that the excitation gap $\Delta = \left(1 - \cos\frac{2\pi}{3}\right) J$ exceeds the cutoff Ω. Then

$$\begin{aligned} P^{(\Omega)} &= \sum_{\alpha=-1}^{+1} |\psi_{k_\alpha}\rangle\, \theta(\Omega - \epsilon_{k_\alpha} + \epsilon_{k_0})\, \langle\psi_{k_\alpha}| \\ &= |\psi_{k_0}\rangle\, \langle\psi_{k_0}| \end{aligned} \tag{70}$$

is simply a projector on the ground state. In this case

$$P_n^{(\Omega)} = |\psi_{k_0}\rangle\, \frac{1}{3}\, \langle\psi_{k_0}| \tag{71}$$

and

$$\left\langle \psi_{k_0} \left| [P_n^{(\Omega)}]^2 \right| \psi_{k_0} \right\rangle - \left\langle \psi_{k_0} | P_n^{(\Omega)} | \psi_{k_0} \right\rangle^2 = 0. \tag{72}$$

The charge fluctuations in the ground state are then zero (as measured by the finite bandwidth detector).

The argument for the Laughlin quasiparticles is similar. We again emphasize that one can not think of a single charge tunneling among three sites because the excitation gap remains finite no matter how far apart the quasiparticle sites are located. This is possible only because it is a correlated many-particle system.

To gain a better understanding of fractional charge it is useful to compare this situation to that in high energy physics. In that field of study one knows the physics at low energies — this is just the phenomena of our everyday world. The goal is to study the high energy (short length scale) limit to see where this low energy physics comes from. What force laws lead to our world? Probing the proton with high energy electrons we can temporarily break it up into three fractionally charged quarks, for example.

Condensed matter physics in a sense does the reverse. We know the phenomena at 'high' energies (i.e. room temperature) and we would like to see how the known dynamics (Coulomb's law and non-relativistic quantum mechanics) leads to unknown and surprising collective effects at low temperatures and long length scales. The analog of the particle accelerator is the dilution refrigerator.

To further understand Laughlin quasiparticles consider the point of view of 'flatland' physicists living in the cold, two-dimensional world of a $\nu = 1/3$ quantum Hall sample. As far as the flatlanders are concerned the 'vacuum' (the Laughlin liquid) is completely inert and featureless. They discover however that the universe is not completely empty. There are a few elementary particles around, all having the same charge q. The flatland equivalent of Benjamin Franklin chooses a unit of charge which not only makes q negative but gives it the fractional value $-1/3$. For some reason the Flatlanders go along with this.

Flatland cosmologists theorize that these objects are 'cosmic strings', topological defects left over from the 'big cool down' that followed the creation of the universe. Flatland experimentalists call for the creation of a national accelerator facility which will reach the unprecedented energy scale of 10 Kelvin. With great effort and expense this energy scale is reached and the accelerator is used to smash together three charged particles. To the astonishment of the entire world a new short-lived particle is temporarily created with the bizarre property of having integer charge!

There is another way to see that the Laughlin quasiparticles carry fractional charge which is useful to understand because it shows the deep connection between the sharp fractional charge and the sharp quantization of the Hall conductivity. Imagine piercing the sample with an infinitely thin magnetic solenoid as shown in Fig. (10) and slowly increasing the magnetic flux Φ from 0 to $\Phi_0 = \frac{hc}{e}$ the quantum of flux. Because of the existence of a finite excitation gap Δ the process is adiabatic and reversible if performed slowly on a time scale long compared to $\hbar/\Delta$.

Faraday's law tells us that the changing flux induces an electric field obeying

$$\oint_\Gamma d\vec{r} \cdot \vec{E} = -\frac{1}{c}\, \frac{\partial \Phi}{\partial t} \tag{73}$$

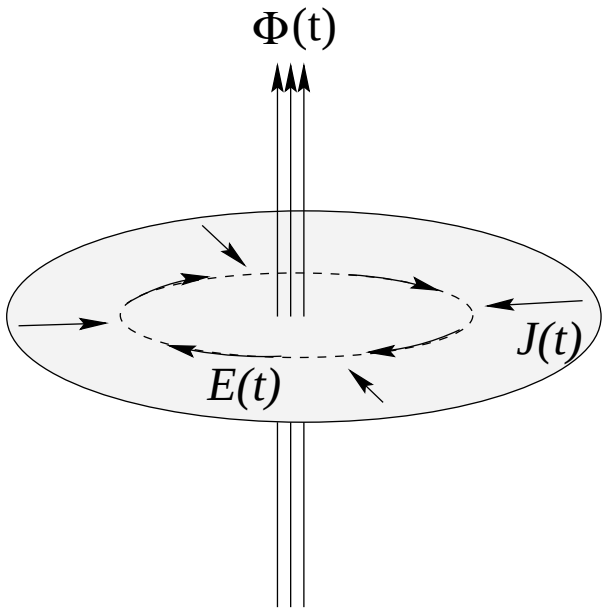

Figure 10: Construction of a Laughlin quasiparticle by adiabatically threading flux $\Phi(t)$ through a point in the sample. Faraday induction gives an azimuthal electric field $E(t)$ which in turn produces a radial current $J(t)$. For each quantum of flux added, charge νe flows into (or out of) the region due to the quantized Hall conductivity $\nu e^2/h$. A flux tube containing an integer number of flux quanta is invisible to the particles (since the Aharanov phase shift is an integer multiple of 2π) and so can be removed by a singular gauge transformation.

where Γ is any contour surrounding the flux tube. Because the electric field contains only Fourier components at frequencies ω obeying $\hbar\omega < \Delta$, there is no dissipation and $\sigma_{xx} = \sigma_{yy} = \rho_{xx} = \rho_{yy} = 0$. The electric field induces a current density obeying

$$\vec{E} = \rho_{xy}\ \vec{J} \times \hat{z} \tag{74}$$

so that

$$\rho_{xy} \oint_\Gamma \vec{J} \cdot (\hat{z} \times d\vec{r}) = -\frac{1}{c}\ \frac{d\Phi}{dt}. \tag{75}$$

The integral on the LHS represents the total current flowing into the region enclosed by the contour. Thus the charge inside this region obeys

$$\rho_{xy}\ \frac{dQ}{dt} = -\frac{1}{c}\ \frac{d\Phi}{dt}. \tag{76}$$

After one quantum of flux has been added the final charge is

$$Q = \frac{1}{c}\ \sigma_{xy}\Phi_0 = \frac{h}{e}\ \sigma_{xy}. \tag{77}$$

Thus on the quantized Hall plateau at filling factor ν where $\sigma_{xy} = \nu\ \frac{e^2}{h}$ we have the result

$$Q = \nu e. \tag{78}$$

Reversing the sign of the added flux would reverse the sign of the charge.

The final step in the argument is to note that an infinitesimal tube containing a quantum of flux is invisible to the particles. This is because the Aharonov-Bohm phase factor for traveling around the flux tube is unity.

$$\exp\left\{ i\frac{e}{\hbar c} \oint_\Gamma \delta\vec{A} \cdot d\vec{r} \right\} = e^{\pm 2\pi i} = 1. \tag{79}$$

Here $\delta\vec{A}$ is the additional vector potential due to the solenoid. Assuming the flux tube is located at the origin and making the gauge choice

$$\delta\vec{A} = \Phi_0 \, \frac{\hat{\theta}}{2\pi r}, \tag{80}$$

one can see by direct substitution into the Schrödinger equation that the only effect of the quantized flux tube is to change the phase of the wave function by

$$\psi \to \psi \prod_j \frac{z_j}{|z_j|} = \psi \prod_j e^{i\theta_j}. \tag{81}$$

The removal of a quantized flux tube is thus a 'singular gauge change' which has no physical effect.

Let us reiterate. Adiabatic insertion of a flux quantum changes the state of the system by pulling in (or pushing out) a (fractionally) quantized amount of charge. Once the flux tube contains a quantum of flux it effectively becomes invisible to the electrons and can be removed by means of a singular gauge transformation.

Because the excitation gap is preserved during the adiabatic addition of the flux, the state of the system is fully specified by the position of the resulting quasiparticle. As discussed before there are no low-lying quasi-degenerate states. This version of the argument highlights the essential importance of the fact that $\sigma_{xx} = 0$ and σ_{xy} is quantized. The existence of the fractionally quantized Hall transport coefficients guarantees the existence of fractionally charged elementary excitations

These fractionally charged objects have been observed directly by using an ultrasensitive electrometer made from a quantum dot [23] and by the reduced shot noise which they produce when they carry current [24].

Because the Laughlin quasiparticles are discrete objects they cost a non-zero (but finite) energy to produce. Since they are charged they can be thermally excited only in neutral pairs. The charge excitation gap is therefore

$$\Delta_c = \Delta_+ + \Delta_- \tag{82}$$

where $\Delta_\pm$ is the vortex/antivortex (quasielectron/quasihole) excitation energy. In the presence of a transport current these thermally excited charges can move under the influence of the Hall electric field and dissipate energy. The resulting resistivity has the Arrhenius form

$$\rho_{xx} \sim \gamma \frac{h}{e^2} \, e^{-\beta\Delta_c/2} \tag{83}$$

where γ is a dimensionless constant of order unity. Note that the law of mass action tells us that the activation energy is $\Delta_c/2$ not Δ_c since the charges are excited in pairs. There is a close analogy between the dissipation described here and the flux flow resistance caused by vortices in a superconducting film.

Theoretical estimates of Δ_c are in good agreement with experimental values determined from transport measurements [25]. Typical values of Δ_c are only a few percent of $e^2/\epsilon\ell$ and hence no larger than a few Kelvin. In a superfluid time-reversal symmetry guarantees that vortices and antivortices have equal energies. The lack of time reversal symmetry here means that Δ_+ and Δ_- can be quite different. Consider for example the hard-core model for which the Laughlin wave function ψ_m is an exact zero energy ground state as shown in eq. (41). Equation (58) shows that the quasihole state contains ψ_m as a factor and hence is also an exact zero energy eigenstate for the hard-core interaction. Thus the quasihole costs zero energy. On the other hand eq. (59) tells us that the derivatives reduce the degree of homogeneity of the Laughlin polynomial and therefore the energy of the quasielectron *must* be non-zero in the hard-core model. At filling factor $\nu = 1/m$ this asymmetry has no particular significance since the quasiparticles must be excited in pairs.

Consider now what happens when the magnetic field is increased slightly or the particle number is decreased slightly so that the filling factor is slightly smaller than $1/m$. The lowest energy way to accommodate this is to inject m quasiholes into the Laughlin state for each electron that is removed (or for each $m\Phi_0$ of flux that is added). The system energy (ignoring disorder and interactions in the dilute gas of quasiparticles) is

$$E_+ = E_m - \delta N \; m\Delta_+ \tag{84}$$

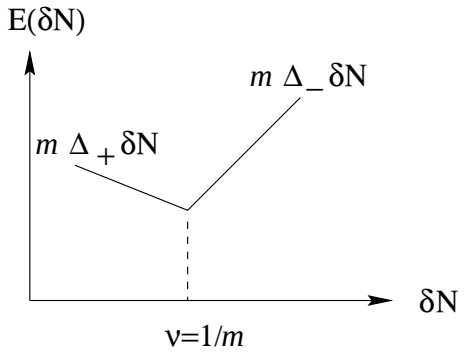

Figure 11: Energy cost for inserting δN electrons into the Laughlin state near filling factor $\nu = 1/m$. The slope of the line is the chemical potential. Its discontinuity at $\nu = 1/m$ measures the charge excitation gap.

where E_m is the Laughlin ground state energy and $-\delta N$ is the number of added holes. Conversely for filling factors slightly greater than $1/m$ the energy is (with $+\delta N$ being the number of added electrons)

$$E_- = E_m + \delta N \; m\Delta_-. \tag{85}$$

This is illustrated in Fig. (11). The slope of the lines in the figure determines the chemical potential

$$\mu_\pm = \frac{\partial E_\pm}{\partial \delta N} = \mp m\Delta_\pm. \tag{86}$$

The chemical potential suffers a jump discontinuity of $m(\Delta_+ + \Delta_-) = m\Delta_c$ just at filling factor $\mu = 1/m$. This jump in the chemical potential is the signature of the charge excitation gap just as it is in a semiconductor or insulator. Notice that this form of the energy is very reminiscent of the energy of a type-II superconductor as a function of the applied magnetic field (which induces vortices and therefore has an energy cost $\Delta E \sim |B|$).

Recall that in order to have a quantized Hall plateau of finite width it is necessary to have disorder present. For the integer case we found that disorder localizes the excess electrons allowing the transport coefficients to not change with the filling factor. Here it is the fractionally-charged quasiparticles that are localized by the disorder.[6] Just as in the integer case the disorder may fill in the gap in the density of states but the DC value of σ_{xx} can remain zero because of the localization. Thus the fractional plateaus can have finite width.

If the density of quasiparticles becomes too high they may delocalize and condense into a correlated Laughlin state of their own. This gives rise to a hierarchical family of Hall plateaus at rational fractional filling factors $\nu = p/q$ (generically with q odd due to the Pauli principle). There are several different but entirely equivalent ways of constructing and viewing this hierarchy which we will not delve into here [2, 3, 5].

5 Summary

In these notes I have discussed the Laughlin ground state and the basic facts of the neutral and charged collective excitations above it. These topics barely scratch the surface of the rich phenomenology of two-dimensional electron gases in the quantum Hall regime. The reader interested in further details about fractional statistics, edge states, Chern-Simons field theories, bilayer quantum Hall systems, quantum Hall ferromagnets and other more advanced topics is directed to the various reviews that are available. [2–11]

[6] Note again the essential importance of the fact that the objects are 'elementary particles'. That is, there are no residual degeneracies once the positions are pinned down.

6 Acknowledgments

Much of my work on the quantum Hall effect has been in collaboration with my long time friend and collaborator Allan MacDonald. This work was supported by National Science Foundation Grant DMR-0342157.

References

[1] H. L. Störmer, *Physica* **B177**, 401 (1992).

[2] *The Quantum Hall Effect, 2nd Ed.*, edited by Richard E. Prange and Steven M. Girvin (Springer-Verlag, New York, 1990).

[3] T. Chakraborty and P. Pietiläinen, *The Fractional Quantum Hall Effect* (Springer-Verlag, Berlin, New York, 1988).

[4] Allan H. MacDonald, *Quantum Hall Effect: A Perspective* (Kluwer Academic Publishers, 1989).

[5] *Perspectives in Quantum Hall Effects*, Edited by Sankar Das Sarma and Aron Pinczuk (Wiley, New York, 1997).

[6] *Introduction to the Theory of the Integer Quantum Hall Effect*, M. Janßen, O. Viehweger, U. Fastenrath, and J. Hajdu (VCH, Weinheim, New York, 1994).

[7] *Quantum Hall Effect*, edited by Michael Stone (World Scientific, Singapore, 1992).

[8] S. Kivelson, D.-H. Lee and S.-C. Zhang, *Scientific American*, March, 1996, p. 86.

[9] Shou Cheng Zhang, *Int. J. Mod. Phys.* B**6**, 25 (1992).

[10] A. H. MacDonald, in *Mesoscopic Quantum Physics*, Les Houches, Session LXI, eds. E. Akkermans, G. Montambaux, J.-L. Pichard and J. Zinn-Justin (North Holland, Amsterdam, 1995).

[11] S. M. Girvin, 'The Quantum Hall Effect: Novel Excitations and Broken Symmetries,' 120 pp. Les Houches Lecture Notes, in: *Topological Aspects of Low Dimensional Systems*, ed. by Alain Comtet, Thierry Jolicoeur, Stephane Ouvry and Francois David, (Springer-Verlag, Berlin and Les Editions de Physique, Les Ulis, 2000, ISBN: 3-540-66909-4), (eprint: cond-mat/9907002).

[12] S. M. Girvin and T. Jach, *Phys. Rev.* B**29**, 5617 (1984).

[13] V. Bargman, *Rev. Mod. Phys.* **34**, 829 (1962).

[14] D. Levesque, J. J. Weiss, and A. H. MacDonald, *Phys. Rev.* B**30**, 1056 (1984).

[15] R. P. Feynman, *Statistical Mechanics* (Benjamin, Reading, 1972).

[16] S. M. Girvin, A. H. MacDonald and P. M. Platzman, *Phys. Rev.* B**33**, 2481 (1986).

[17] S. M. Girvin in Chap. 10 and App. I of Ref. [2]; S. M. Girvin and A. H. MacDonald, *Phys. Rev. Lett.* **58**, 1252 (1987); S.-C. Zhang, H. Hansson, and S. Kivelson, *Phys. Rev. Lett.* **62**, 82 (1989); N. Read, *Phys. Rev. Lett.* **62**, 86 (1989); D.-H. Lee and M. P. A. Fisher, *Phys. Rev. Lett.* **63**, 903 (1989).

[18] F. D. M. Haldane and E. H. Rezayi, *Phys. Rev. Lett.* **54**, 237 (1985).

[19] G. Fano, F. Ortolani, and E. Colombo, *Phys. Rev.* B**34**, 2670 (1986).

[20] A. Pinczuk, B. S. Dennis, L. N. Pfeiffer, and K. W. West, *Phys. Rev. Lett.* **70**, 3983 (1993).

[21] C. Kallin and B. I. Halperin, *Phys. Rev.* B**30**, 5655 (1984); *Phys. Rev.* B**31**, 3635 (1985).

[22] A. Goldhaber and S. A. Kivelson, *Physics Lett.* B**255**, 445 (1991).

[23] V. Goldman and B. Su, *Science* **267**, 1010 (1995).

[24] R. de-Picciotto, M. Reznikov, M. Heiblum, V. Umansky, G. Bunin, and D. Mahalu, *Nature* **389**, 162 (1997); L. Saminadayar, D. C. Glattli, Y. Jin and B. Etienne, *Phys. Rev. Lett.* **79**, 2526 (1997).

[25] R. L. Willett, H. L. Störmer, D. C. Tsui, A. C. Gossard, and J. H. English, *Phys. Rev.* B**37**, 8476 (1988).