Lecture Notes on Quantum Hall Effect
(A Work in Progress)

Daniel Arovas
Department of Physics
University of California, San Diego

September 3, 2020

Contents

List of Figures

List of Tables

0.1 Preface

This is a proto-preface. A more complete preface will be written after these notes are completed. Which may not be for some time.

These notes are dedicated to the memory of my dog, Henry.

Figure 1: Henry (2000 - 2015)

Chapter 1

Preliminaries

1.1 Introduction

The quantum Hall effect (QHE) refers to a set of phenomena and associated phases of matter found in two-dimensional electron gases subjected to a large perpendicular magnetic field[1]. The phenomena are typically divided into two classes, the *integer quantum Hall effect* (IQHE) and the *fractional quantum Hall effect* (FQHE), depending on the *Landau level filling fraction*, given by $\nu = nhc/eB$, where n is the two-dimensional electron density and B the magnetic field strength. The combination $\phi_0 = hc/e = 4.137 \times 10^{-7}\,\mathrm{G}\cdot\mathrm{cm}^2$ is the Dirac flux quantum[2], hence

$$\nu = 4.14 \cdot n[10^{11}\,\mathrm{cm}^{-2}]/B[\mathrm{T}] \quad . \tag{1.1}$$

Thus, in a field of $B = 4.14\,\mathrm{T}$, the Landau level (LL) filling fraction $\nu = 1$ occurs for an electron density $n = 10^{11}\,\mathrm{cm}^{-2}$.

The IQHE was discovered by von Klitzing in 1980 in routine magnetotransport studies of silicon MOSFETs[3]. The FQHE was discovered by Tsui and Störmer in 1982 [4], in GaAs$-$Al$_x$Ga$_{1-x}$As heterojunctions. The experimental setup is depicted in Fig. 1.1, and some spectacular data shown in Fig. 1.2. An electrical current I is established along the $\hat{\boldsymbol{x}}$ direction, and the longitudinal and transverse voltage drops V_{L} and V_{H} are measured, from which one obtains, in the linear response regime, the resistances $R_{\mathrm{L}} = V_{\mathrm{L}}/I$ and $R_{\mathrm{H}} = V_{\mathrm{H}}/I$. In the IQHE, one observes that R_{H} remains constant along plateaus as the filling fraction ν is varied (either by varying the electron density n, typically with a gate, or by varying the magnetic field B). The plateau values are given by $R_{\mathrm{H}} = h/pe^2$, where $p \in \mathbb{Z}$ is an integer, for $\nu = p$. In the FQHE, one observes

[1] The effect has been seen in hole gases as well.

[2] This is often more conveniently expressed as $\phi_0 = 4.137 \times 10^5\,\mathrm{T}\cdot\mathrm{\AA}^2$, where $1\,\mathrm{T} = 10^4\,\mathrm{G}$.

[3] K. von Klitzing, G. Dorda, and M. Pepper, *Phys. Rev. Lett.* **45**, 494 (1980).

[4] D. C. Tsui, H. L. Störmer, and A. C. Gossard, *Phys. Rev. Lett.* **48**, 1559 (1982).

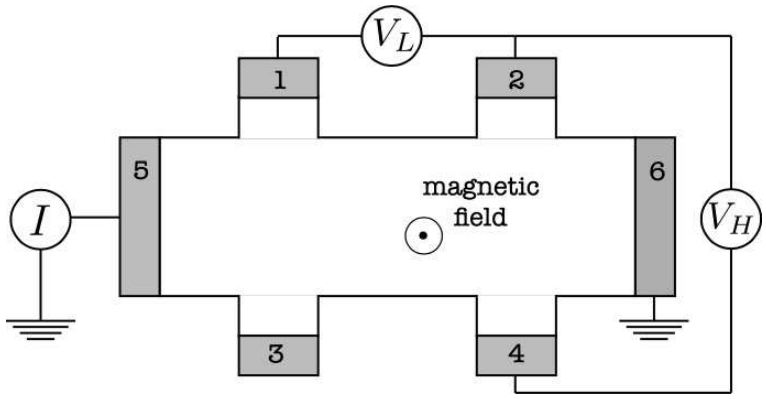


Figure 1.1: A Hall bar setup.

plateaus at rational fractions $\nu = p/q$, typically with q odd[5], where $R_{\rm H} = qh/pe^2$. The quantity $R_{\rm Q} = h/e^2 = 25,812.8\,\Omega$ is known as the *quantum of resistance*.

1.1.1 Resistance, conductance, resistivity, conductivity

In the linear response regime, one has $V_\alpha = R_{\alpha\beta}\, I_\beta$, where R is the resistance tensor. It's matrix inverse, $G = R^{-1}$, is known as the conductance, with $I_\alpha = G_{\alpha\beta}\, V_\beta$. The units of each element $R_{\alpha\beta}$ of the resistance tensor are Ohms (Ω), hence the units of $G_{\alpha\beta}$ are Ω^{-1}.

Resistance and conductance are not materials parameters (*i.e.* intensive quantities); you can't look up the resistance of copper in a table, for example. If, *ceteris paribus*, you double the length of a copper wire, its resistance doubles[6]. What doesn't change is the metal's *resistivity*, ρ, which *is* a materials parameter[7]. The corresponding linear response relation is between *current density* and *electric field*, *viz.* $E_\alpha = \rho_{\alpha\beta}\, j_\beta$. The inverse of the resistivity tensor is the conductivity tensor $\sigma = \rho^{-1}$, for which $j_\alpha = \sigma_{\alpha\beta} E_\beta$.

For an isotropic d-dimensional cube of side length L, in zero magnetic field, if the current along one of the cubic axes is I then the current density is $j = I/L^{d-1}$. Similarly, if the voltage drop along this axis is V, the electric field is $E = V/L$. Thus $R = V/I = \rho L^{2-d}$, and we see that resistance and resistivity in general have different units. Similarly $G = \sigma L^{d-2}$. In two dimensions, resistance and resistivity have the same dimensions, but nevertheless resistance is a *geometric* quantity. Consider a $L_x \times L_y$ rectangular sample with conductivity tensor

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ \sigma_{yx} & \sigma_{yy} \end{pmatrix} \quad , \tag{1.2}$$

[5]The even denominator quantum Hall effect is very interesting and distinct from the odd denominator effect.

[6]Assuming, that is, that the length L is longer than the inelastic scattering (or *phase breaking*) length, ℓ_ϕ. For $L < \ell_\phi$, quantum interference effects become important and Ohm's law is no longer valid.

[7]The resistivity will in general depend on the temperature, and on the density and type of impurities present, as well as on the material itself.

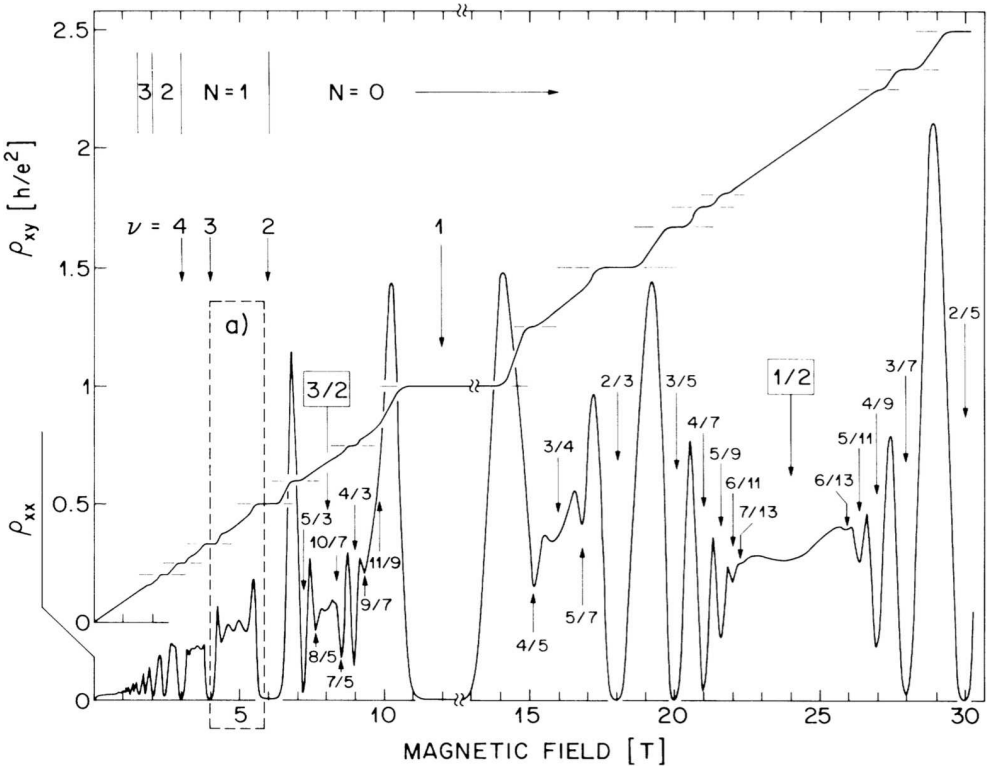


Figure 1.2: Low temperature ($T \approx 150\,\mathrm{mK}$) longitudinal resistivity ρ_{xx} and Hall resistivity ρ_{xy} as a function of applied magnetic field in a two-dimensional electron gas system (GaAs/AlGaAs heterostructure), from R. Willett *et al.*, *Phys. Rev. Lett.* **59**, 1776 (1987). Each dip in ρ_{xx} and concomitant plateau in ρ_{xy} corresponds to a distinct phase of matter.

with $\boldsymbol{j} = \sigma \boldsymbol{E}$. In general, linear response transport is described by the set of equations $J_i = L_{ik} F_k$, where the $\{J_i\}$ are generalized currents and the $\{F_k\}$ generalized forces. Onsager reciprocity[8] then requires

$$L_{ik}(\boldsymbol{B}) = \eta_i \, \eta_k \, L_{ki}(-\boldsymbol{B}) \quad , \tag{1.3}$$

with no sum on i or k, where $\eta_i = \pm 1$ according to whether J_i is symmetric or antisymmetric under time reversal, *i.e.* $J_i^{\mathcal{T}} = \eta_i J_i$. Thus, $\sigma_{yx}(\boldsymbol{B}) = \sigma_{xy}(-\boldsymbol{B})$ since both j_x and j_y are odd under time reversal. But $\boldsymbol{B} \to -\boldsymbol{B}$ reverses the orientation of the $(\hat{\boldsymbol{x}}, \hat{\boldsymbol{y}}, \hat{\boldsymbol{B}})$ triad, hence $\sigma_{xy}(-\boldsymbol{B}) = -\sigma_{xy}(\boldsymbol{B})$, and we have that the off-diagonal elements of the conductivity tensor are antisymmetric: $\sigma_{yx}(\boldsymbol{B}) = -\sigma_{xy}(\boldsymbol{B})$. Now let's write the current densities as $j_x = I_x / L_y$ and $j_y = I_y / L_x$, and the fields as $E_x = V_x / L_x$ and $E_y = V_y / L_y$. We then have

$$\overbrace{\begin{pmatrix} L_y^{-1} & 0 \\ 0 & L_x^{-1} \end{pmatrix} \begin{pmatrix} I_x \\ I_y \end{pmatrix}}^{\boldsymbol{j}} = \overbrace{\begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{yy} \end{pmatrix}}^{\sigma} \overbrace{\begin{pmatrix} L_x^{-1} & 0 \\ 0 & L_y^{-1} \end{pmatrix} \begin{pmatrix} V_x \\ V_y \end{pmatrix}}^{\boldsymbol{E}} \quad , \tag{1.4}$$

[8] See L. D. Landau and E. M. Lifshitz, *Statistical Physics*, part I, §120.

from which we read off the relation between conductance and conductivity tensors,

$$\begin{pmatrix} G_{xx} & G_{xy} \\ G_{yx} & G_{yy} \end{pmatrix} = \begin{pmatrix} L_y & 0 \\ 0 & L_x \end{pmatrix} \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{yy} \end{pmatrix} \begin{pmatrix} L_x^{-1} & 0 \\ 0 & L_y^{-1} \end{pmatrix} = \begin{pmatrix} \frac{L_y}{L_x}\,\sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \frac{L_x}{L_y}\,\sigma_{yy} \end{pmatrix} \quad . \tag{1.5}$$

Similarly, the relation between resistance and resistivity tensors is

$$\begin{pmatrix} R_{xx} & R_{xy} \\ R_{yx} & R_{yy} \end{pmatrix} = \begin{pmatrix} \frac{L_x}{L_y}\,\rho_{xx} & \rho_{xy} \\ -\rho_{xy} & \frac{L_y}{L_x}\,\rho_{yy} \end{pmatrix} \quad . \tag{1.6}$$

Finally,

$$\rho = \begin{pmatrix} \rho_{xx} & \rho_{xy} \\ -\rho_{xy} & \rho_{yy} \end{pmatrix} = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{yy} \end{pmatrix}^{-1} = \frac{1}{\sigma_{xx}^2 + \sigma_{xy}^2} \begin{pmatrix} \sigma_{yy} & -\sigma_{xy} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix} = \sigma^{-1} \quad . \tag{1.7}$$

Along the QH plateaus, as $T \to 0$, the longitudinal resistivity vanishes as $\rho_{xx}(T) \propto e^{-\Delta/k_{\rm B}T}$, where Δ is the energy gap for transport. Thus, at $T = 0$ the resistivity and conductivity tensors are purely off-diagonal, with $\rho_{\alpha\beta} = \rho_{xy}\,\epsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta} = \sigma_{xy}\,\epsilon_{\alpha\beta}$, with $\rho_{xy} = 1/\sigma_{xy}$.

1.1.2 Semiclassical magnetotransport theory

Combining Newton's second law with the Lorentz force law for a particle of charge $-e$ and mass m, we have

$$\frac{d\boldsymbol{p}}{dt} = -e\boldsymbol{E} - \frac{e}{c}\,\frac{\boldsymbol{p}}{m} \times \boldsymbol{B} - \frac{\boldsymbol{p}}{\tau} \quad , \tag{1.8}$$

where the last term is a frictional force which in metals and semiconductors typically comes from electron-impurity scattering[9], with τ the transport scattering time[10]. We take $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, and write the current density as $\boldsymbol{j} = -ne\boldsymbol{p}/m$. Defining the *cyclotron frequency* $\omega_{\rm c} = eB/mc$, and setting $\dot{\boldsymbol{p}} = 0$ in steady state, we obtain

$$\frac{ne^2\tau}{m}\,\boldsymbol{E} + \omega_{\rm c}\tau\,\boldsymbol{j} \times \hat{\boldsymbol{z}} + \boldsymbol{j} = 0 \quad , \tag{1.9}$$

the solution of which is $\boldsymbol{j} = \sigma\boldsymbol{E}$, where the conductivity tensor is

$$\sigma = \frac{ne^2\tau/m}{1 + \omega_{\rm c}^2\tau^2} \begin{pmatrix} 1 & -\omega_{\rm c}\tau \\ \omega_{\rm c}\tau & 1 \end{pmatrix} \quad . \tag{1.10}$$

Taking the inverse, we have $\boldsymbol{E} = \rho\boldsymbol{j}$, with resistivity tensor

$$\rho = \sigma^{-1} = \frac{m}{ne^2\tau} \begin{pmatrix} 1 & \omega_{\rm c}\tau \\ -\omega_{\rm c}\tau & 1 \end{pmatrix} \quad . \tag{1.11}$$

[9] Electron-phonon scattering, electron-electron scattering, and boundary scattering are also present.

[10] There is an important difference between the single particle scattering time $\tau_{\rm sp}$ and the transport scattering time $\tau_{\rm tr}$. See, *e.g.*, §1.5 of my Physics 211B lecture notes for details.

What is n? Naïvely one might think it is the total electron density, but of course this is wrong. As we know from elementary solid state physics, filled Bloch bands are inert and carry no net current. A somewhat more realistic linearized Boltzmann equation approach, assuming an isotropic parabolic conduction band with electron carriers, yields the same result, with $n = \int d\varepsilon\, g_{\rm c}(\varepsilon)\, f^0(\varepsilon - \mu)$ the conduction electron density, with $g_{\rm c}(\varepsilon)$ is the conduction band density of states and $f^0(\varepsilon - \mu)$ the Fermi function, and m replaced by the effective mass m^* of the conduction band. All the fully occupied bands below the conduction band contribute nothing to the current. Note that $\rho_{xx} = \rho_{yy} = m^*/ne^2\tau$ because the system is isotropic. For the anisotropic parabolic band, where the effective mass tensor $m^*_{\alpha\beta}$ has eigenvalues $m^*_{x,y}$, then along its principal axes one of course has $\rho_{xx} = m^*_x/ne^2\tau$ and $\rho_{yy} = m^*_y/ne^2\tau$, with $\rho_{xy} = B/nec$ as in the isotropic case.

One interesting feature of the semiclassical Boltzmann result is that the diagonal terms of the resistivity tensor are independent of magnetic field. Thus, $\partial\rho_{xx}/\partial B = 0$, and the *magnetoresistance* $\Delta\rho_{xx}(B) \equiv \rho_{xx}(B) - \rho_{xx}(0)$ vanishes. This is in general not the case if one has multiple bands contributing to the transport current (say conduction electrons as well as valence holes), or in the case where the Fermi surface has open orbits which span the Brillouin zone. Thus, as a function of B, the semiclassical result says that $\rho_{xx}(B)$ is constant and $\rho_{xy}(B)$ is perfectly linear. This it completely different from the results shown in Fig. 1.2, except in the very low field regime.

1.1.3 Mobility, cyclotron frequency, and electron-electron interactions

The *mobility* μ is defined by the combination $\mu = e\tau/m^*$. Thus, in zero field, the steady state electron velocity is $\boldsymbol{v} = \mu\boldsymbol{E}$, so mobility has units of $[\mu] = \mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$. In MOSFETs, mobilities are seldom more than a few tens of thousands in these units. But in MBE-grown GaAs heterostructures, mobilities as high as $10^7\,\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$ have been achieved. In GaAs, where the conduction band is isotropic and has effective mass $m^* = 0.067\,m_{\rm e}$, one finds

$$\tau = 3.8 \times 10^{-17}\mathrm{s}\cdot\mu\big[\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}\big] \quad . \tag{1.12}$$

Thus, for $\mu = 10^6\,\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$, one obtains $\tau \simeq 38\,\mathrm{ps}$.

The *cyclotron frequency* is given by the combination $\omega_{\rm c} = eB/m^*c$. With

$$\phi_0 = \frac{hc}{e} = 4.14{\times}10^{-7}\mathrm{G}\cdot\mathrm{cm}^2 \quad , \quad h = 6.63{\times}10^{-27}\mathrm{erg}\cdot\mathrm{s} = 4.14{\times}10^{-15}\mathrm{eV}\cdot\mathrm{s} \quad , \quad k_{\rm B} = 8.62{\times}10^{-5}\mathrm{eV/K} \quad , \tag{1.13}$$

Thus, for GaAs conduction electrons, one obtains

$$\omega_{\rm c} = 2.63{\times}10^{12}\,\mathrm{Hz}B[\mathrm{T}] \quad , \quad \omega_{\rm c}\tau = 10^{-4}\,\mu[\mathrm{cm}^2/\mathrm{Vs}]\,B[\mathrm{T}] \quad , \quad \hbar\omega_{\rm c} = 1.73\,\mathrm{meV}B[T] = 20\,\mathrm{K}\,k_{\rm B}\,B[T] \quad . \tag{1.14}$$

At fields $B \sim 10\,\mathrm{T}$ and in samples of mobility $\mu \sim 10^6\,\mathrm{cm}^2/\mathrm{V}\cdot\mathrm{s}$, we have $\omega_{\rm c}\tau \sim 1000 \gg 1$.

As we shall see, quantization introduces a new length scale, $\ell = (\hbar c/eB)^{1/2}$, called the *magnetic length*. This depends only on physical constants and the magnetic field strength. One finds

$$\ell = (\phi_0/2\pi B)^{1/2} = 257\,\text{Å}/\sqrt{B[T]} \quad . \tag{1.15}$$

From this length scale, we construct the energy scale $e^2/\epsilon\ell$ for electron-electron interactions. For GaAs, where $\epsilon = 13$, we have

$$\frac{e^2}{\epsilon\ell} = 4.31\,\text{meV}\cdot\sqrt{B[T]} = 50.0\,\text{K}\,k_{\text{B}}\sqrt{B[T]} \quad . \tag{1.16}$$

1.1.4 $\vec{E}\times\vec{B}$ drift and separation of time scales

For a classical particle of charge e moving in the (x,y) plane and subjected to a magnetic field $\boldsymbol{B} = B\hat{\boldsymbol{z}}$, the equations of motion are given by the Lorentz force law,

$$m\ddot{\boldsymbol{r}} = -\boldsymbol{\nabla}V - \frac{e}{c}\,B\,\dot{\boldsymbol{r}}\times\hat{\boldsymbol{z}} \quad . \tag{1.17}$$

We now write $\boldsymbol{r}(t) = \boldsymbol{\mathcal{R}}(t) + \boldsymbol{\xi}(t)$. We presume that the guiding-center motion $\boldsymbol{\mathcal{R}}(t)$ executes large excursions, slowly drifting along equipotentials of $V(\boldsymbol{r})$, while the cyclotron motion $\boldsymbol{\xi}(t)$ executes fast small excursions with characteristic time scale $2\pi/\omega_{\text{c}}$. This assumption will be borne out in the following analysis.

The zeroth order theory is simply given by

$$\begin{aligned}\dot{\boldsymbol{\mathcal{R}}} &= -\frac{c}{eB}\,\hat{\boldsymbol{z}}\times\boldsymbol{\nabla}V(\boldsymbol{\mathcal{R}}) \\ \dot{\boldsymbol{\xi}} &= \omega_{\text{c}}\,\hat{\boldsymbol{z}}\times\boldsymbol{\xi} \quad .\end{aligned} \tag{1.18}$$

Thus, the guiding-center executes a slow drift in the direction of $\boldsymbol{\nabla}V\times\hat{\boldsymbol{z}}$, while the cyclotron coordinate executes counterclockwise circular motion as viewed from above.

Proceeding with the expansion in powers of the cyclotron motion, we have

$$\begin{aligned}m\ddot{\mathcal{R}}_\alpha + m\ddot{\xi}_\alpha = &-\partial_\alpha V(\boldsymbol{\mathcal{R}}) - \xi_\beta\,\partial_\alpha\,\partial_\beta V(\boldsymbol{\mathcal{R}}) - \tfrac{1}{2}\,\xi_\beta\,\xi_\gamma\,\partial_\alpha\,\partial_\beta\,\partial_\gamma V(\boldsymbol{\mathcal{R}}) + \ldots \\ &- \frac{eB}{c}\,\epsilon_{\alpha\beta}\,\dot{\mathcal{R}}_\beta - \frac{eB}{c}\,\epsilon_{\alpha\beta}\,\dot{\xi}_\beta \quad .\end{aligned} \tag{1.19}$$

Here we have used the relation, for any vector $\boldsymbol{u}$,

$$\epsilon_{\alpha\beta}\,u_\beta = (u_y\,,\,-u_x) = (\boldsymbol{u}\times\hat{\boldsymbol{z}})_\alpha \quad . \tag{1.20}$$

We assume $\dot{\mathcal{R}}_\alpha = 0$ on average, leading to the slow equation,

$$\frac{eB}{c}\,\epsilon_{\alpha\beta}\,\dot{\mathcal{R}}_\beta = -\partial_\alpha V(\boldsymbol{\mathcal{R}}) - \tfrac{1}{2}\left\langle\xi_\beta\,\xi_\gamma\right\rangle\partial_\alpha\,\partial_\beta\,\partial_\gamma V(\boldsymbol{\mathcal{R}}) - \ldots \quad , \tag{1.21}$$

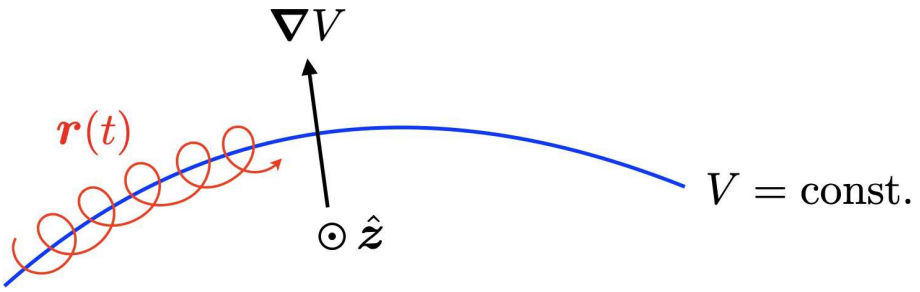


Figure 1.3: Cyclotron motion and guiding-center drift.

where $\langle \xi_\beta \, \xi_\gamma \rangle$ is averaged over the fast motion, and the fast equation,

$$m\ddot{\xi}_\alpha = -\xi_\beta \, \partial_\alpha \, \partial_\beta \, V(\boldsymbol{\mathcal{R}}) - \frac{eB}{c} \, \epsilon_{\alpha\beta} \, \dot{\xi}_\beta + \ldots \quad . \tag{1.22}$$

On the fast scale of the $\boldsymbol{\xi}(t)$ motion, the guiding-center $\boldsymbol{\mathcal{R}}(t)$ is assumed constant. Fourier transforming the fast motion, we write $\boldsymbol{\xi}(t) = \mathsf{Re}\, \boldsymbol{\xi}^0 \, e^{-i\omega t}$, with

$$\begin{pmatrix} -m\omega^2 + V_{xx} & im\omega\omega_{\rm c} + V_{xy} \\ -im\omega\omega_{\rm c} + V_{xy} & -m\omega^2 + V_{yy} \end{pmatrix} \begin{pmatrix} \xi^0_x \\ \xi^0_y \end{pmatrix} = 0 \quad , \tag{1.23}$$

where $V_{\alpha\beta} \equiv \partial_\alpha \, \partial_\beta \, V(\boldsymbol{\mathcal{R}})$. Solving for ω, we take the fast root of the resulting quadratic equation and obtain

$$\begin{aligned} \omega_+^2 &= \frac{1}{2}\left(\omega_{\rm c}^2 + \frac{V_{xx}+V_{yy}}{m}\right) + \frac{1}{2}\,\omega_{\rm c}^2 \sqrt{1 + \frac{2(V_{xx}+V_{yy})}{m\omega_{\rm c}^2} + \frac{(V_{xx}-V_{yy})^2}{m^2\omega_{\rm c}^4} + \frac{V_{xy}^2}{m^2\omega_{\rm c}^4}} \\ &= \omega_{\rm c}^2 + \frac{\nabla^2 V}{m} + \ldots \quad . \end{aligned} \tag{1.24}$$

Thus the local cyclotron frequency is given by $\omega_{\rm c}(\boldsymbol{\mathcal{R}}) = \omega_{\rm c} + \frac{\nabla^2 V(\boldsymbol{\mathcal{R}})}{2m\omega_{\rm c}}$ to lowest nontrivial order.

We will need the corresponding eigenvector for the high frequency root. Writing $\boldsymbol{\xi}^0 \equiv (u\,\xi_0, v\,\xi_0)$, with $|u|^2 + |v|^2 = 1$, we have

$$\begin{aligned} u &= \frac{V_{xy} + im\omega_{\rm c}\omega_+}{\sqrt{(V_{xx} - m\omega_+^2)^2 + |V_{xy} + im\omega_{\rm c}\omega_+|^2}} \\ v &= -\frac{V_{xx} - m\omega_+^2}{\sqrt{(V_{xx} - m\omega_+^2)^2 + |V_{xy} + im\omega_{\rm c}\omega_+|^2}} \quad . \end{aligned} \tag{1.25}$$

Averaging over the cyclotron motion, we find

$$\langle \xi_\alpha \, \xi_\beta \rangle = \tfrac{1}{2}\,\xi_0^2 \begin{pmatrix} |u|^2 & \mathsf{Re}\,(u\bar{v}) \\ \mathsf{Re}\,(u\bar{v}) & |v|^2 \end{pmatrix} \quad . \tag{1.26}$$

Since $\omega_+ \approx \omega_{\rm c}$, we obtain $u \approx \frac{i}{\sqrt{2}}$ and $v \approx \frac{1}{\sqrt{2}}$. Thus the guiding-center motion is given by the equation

$$\frac{eB}{c}\,\epsilon_{\alpha\beta}\,\dot{\mathcal{R}}_\beta = -\partial_\alpha V - \tfrac{1}{4}\,\xi_0^2\,\partial_\alpha\,\nabla^2 V \equiv -\partial_\alpha V_{\rm eff}(\boldsymbol{\mathcal{R}}) \quad , \tag{1.27}$$

where the effective guiding-center potential is

$$V_{\rm eff}(\boldsymbol{\mathcal{R}}) = V(\boldsymbol{\mathcal{R}}) + \tfrac{1}{4}\,\xi_0^2\,\nabla^2 V(\boldsymbol{\mathcal{R}}) \quad . \tag{1.28}$$

This makes good physical sense: as the electron moves slowly along the equipotential, it samples, through its small and fast cyclotron excursions, the local environment, inducing a gradient squared correction to the local value of $V(\boldsymbol{\mathcal{R}})$.

For a classical electron moving in a circular orbit of radius r, setting the centrifugal force $F_{\rm c} = mv^2/r$ equal to the Lorentz force evB/c yields the relation $v = eBR/mc$. The kinetic energy is then $T = \frac{1}{2}mv^2 = e^2B^2r^2/2mc^2$. If we now quantize semiclassically, demanding $\pi r^2 \cdot B = (n+\frac{1}{2})\phi_0$, then $r_n^2 = (2n+1)\ell^2$ where $\ell = (\hbar c/eB)^{1/2}$ is the magnetic length. The kinetic energy is then $T = (n+\frac{1}{2})\hbar\omega_{\rm c}$. Thus $\xi_0 = r_n$ in our above derivation of the effective potential, with n the Landau level index.

The potential $V(\boldsymbol{r})$ is due to extrinsic disorder, arising typically from the irregular placement of the dopant atoms in a heterostructure, or semiconductor-oxide interface disorder in a MOSFET. In heterostructures, the dopant ions are typically several hundred Ångstroms removed from the 2DEG layer, and $V(\boldsymbol{r})$ is smooth on this length scale. Suppose the two dimensional electron gas lies in the plane $z = 0$ and consider a 'delta doping' profile in which the donor density is $N_{\rm d}(x,y,z) = N_{\rm d}(\boldsymbol{r})\,\delta(z-d)$ where d is the distance between the 2DEG and the dopant layer. The electrical potential $\phi(\boldsymbol{r})$ at $\boldsymbol{r} = (x,y)$ in the 2DEG plane is then given by

$$\phi(\boldsymbol{r}) = \int \frac{d^2q}{(2\pi)^2} \int\limits_{-\infty}^{\infty} \frac{dq_z}{2\pi}\,\hat{N}_{\rm d}(\boldsymbol{q})\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\,e^{iq_z d}\,\frac{4\pi e}{\boldsymbol{q}^2+q_z^2} = \int \frac{d^2q}{(2\pi)^2}\,\hat{N}_{\rm d}(\boldsymbol{q})\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}}\,\frac{2\pi e\,\exp(-|\boldsymbol{q}|d)}{|\boldsymbol{q}|} \quad , \tag{1.29}$$

and we see that the components of $\hat{N}_{\rm d}(\boldsymbol{q})$ with high spatial frequency are attenuated exponentially. This smooths out the random potential experienced by the electrons in the 2DEG. MOSFETs are typically much dirtier, with correspondingly lower mobilities, hence $V(\boldsymbol{r})$ there is disorder on shorter length scales. Indeed disorder is *essential* to the quantum Hall effect, since in a pristine system we can always perform a Lorentz boost to a frame where $\boldsymbol{B} = 0$ and deduce $\sigma_{xy} = -nec/B$. (This argument is quite a bit more subtle if there are other features breaking translational symmetry, such as leads and surfaces.)

1.2 MOSFETs and Heterojunctions

Where do two-dimensional electron gases (2DEGs) come from? As noted above, the IQHE was first discovered in silicon MOSFETs while the FQHE was first discovered in GaAs heterostruc-

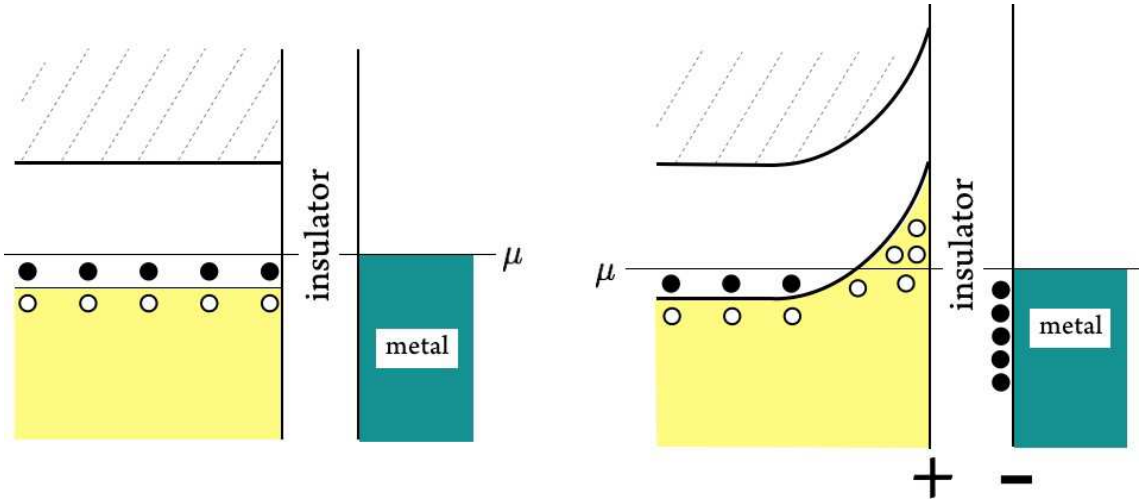


Figure 1.4: Junction between a p-type semiconductor and a metal. Left: Zero bias. Right: Metal biased negative with respect to semiconductor, creating an accumulation layer of holes and a net dipole moment at the interface.

tures. Details of the modeling and important semiconductor physics in these systems are discussed in the 1982 review by Ando, Fowler, and Stern[11]. Today, we have new two-dimensional systems which exhibit the QHE, such as graphene. Graphene is particularly interesting because it is a 'Dirac material' in which the electronic band structure features Dirac points, which are conical intersections of conduction and valence bands described by a two-dimensional Dirac Hamiltonian. More on this later.

In a metal, internal electric fields are efficiently screened and excess charge migrates rapidly to the surface, with charge density fluctuations attenuated exponentially as one enters the bulk. The Thomas-Fermi screening length, $\lambda_{\rm TF} = \big(4\pi e^2 g(\varepsilon_{\rm F})\big)^{-1/2}$, is short (a few Ångstroms) due to the large density of states at the Fermi level. In semiconductors, the Fermi level lies somewhere in the gap between valence and conduction bands, and the density of states at $\varepsilon_{\rm F}$ is quite low. Screening is due to thermally excited charge carriers, and since the carrier density is small in comparison to that in metals, the screening length is many lattice spacings.

Consider now a junction between a semiconductor and a metal, with an intervening insulating layer. This is called MIS, or Metal-Insulator-Semiconductor, junction. If the metal is unbiased relative to the semiconductor, their chemical potentials will align. The situation for a p-type semiconductor - metal junction is depicted in the left panel of Fig. 1.4. Next consider the case in which the metal is biased negatively with respect to the semiconductor, *i.e.* the metal is placed at a negative voltage $-V$. There is then an electric field $\boldsymbol{E} = -\boldsymbol{\nabla}\phi$ pointing *out* of the semiconductor. Electric fields point in the direction positive charges want to move, hence in this case valence holes are attracted to the interface, creating an *accumulation layer* of holes, as

[11] T. Ando, A. B. Fowler, and F. Stern, *Rev. Mod. Phys.* **54**, 437 (1982).

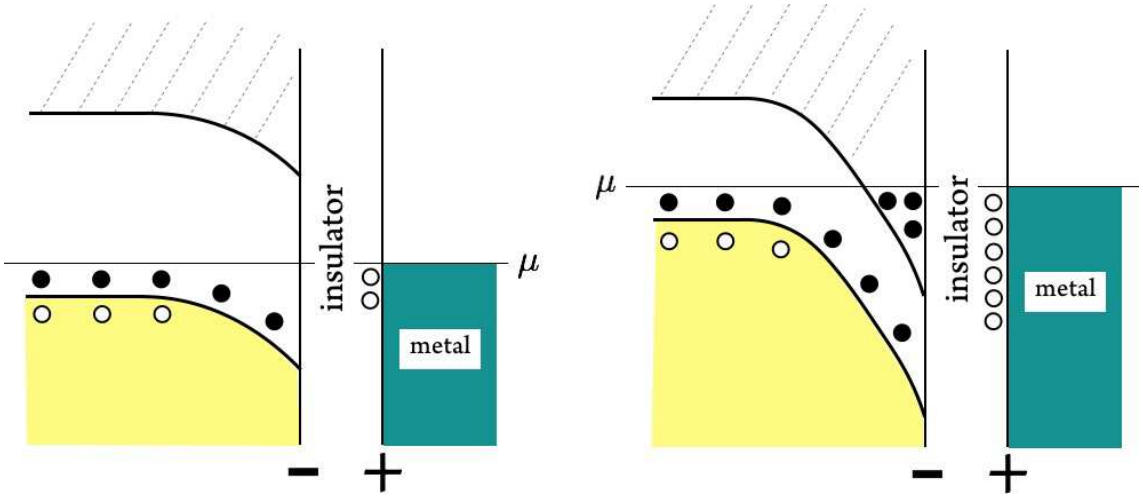


Figure 1.5: Junction between a p-type semiconductor and a metal. Left: Metal biased positive with respect to semiconductor, creating a space charge depletion layer. Right: Strong positive bias creates an inversion layer of n-type carriers in the p-type material.

depicted in the right panel of Fig. 1.4. On the metallic side, electrons migrate to the interface for the same reason. *No charges move across the insulating barrier.* Thus, a dipole layer is created across the barrier, with the dipole moment pointing into the semiconductor. This creates an internal potential whose net difference $\phi_{\rm metal} - \phi_{\rm semiconductor}(-\infty) = V$ exactly cancels the applied bias. This condition in fact is what determines the width of the accumulation layer.

What happens when the metal is biased positively? In this case, the electric field points into the semiconductor, and valence holes are repelled from the semiconductor surface, which is then negatively charged. This, in turn, repels electrons from the nearby metallic surface. The result is a space charge *depletion layer* in the semiconductor, which is devoid of charge carriers (*i.e.* valence holes). This situation is sketched in the left panel of Fig. 1.5.

Finally, what happens if the bias voltage on the metal exceeds the band gap? In this case, the field is so strong that not only are valence holes expelled from the surface, but conduction electrons are present, as shown in the right panel of Fig. 1.5. The presence of n-type carriers in a p-type semiconductor is known as *n-inversion*.

Remember this:

- *Accumulation* : presence of additional n-carriers in an n-type material, or additional p-carriers in a p-type material.
- *Depletion* : absence of n-carriers in an n-type material, or p-carriers in a p-type material.
- *Inversion* : presence of n-carriers in a p-type material, or p-carriers in an n-type material.

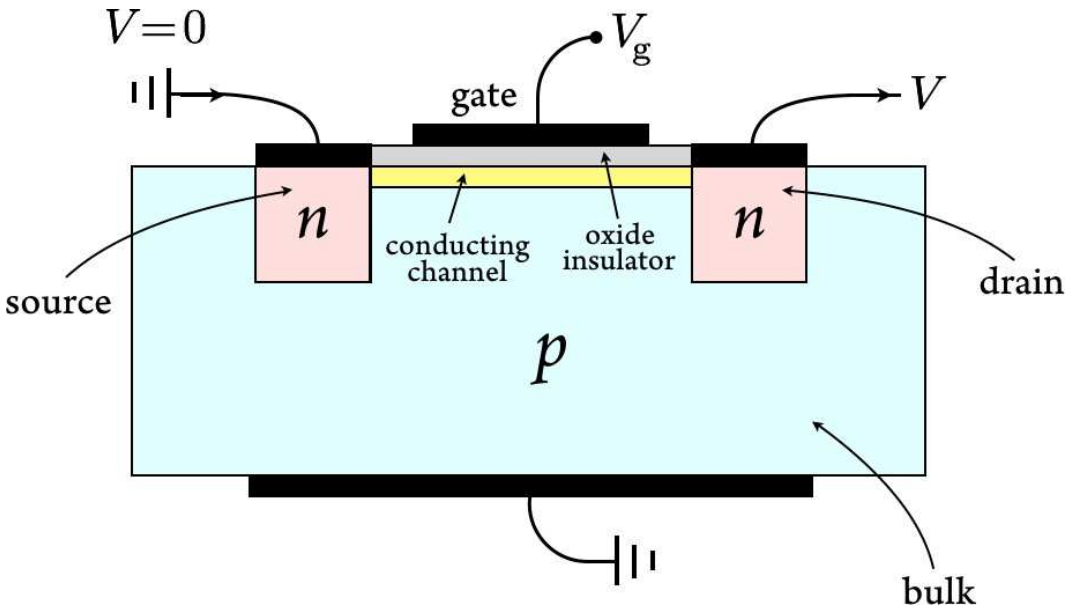


Figure 1.6: The MOSFET.

Inversion occurs when the presence of a depletion layer does not suffice to align the chemical potentials of the two sides of the junction.

1.2.1 The MOSFET

A MOSFET (Metal-Oxide-Semiconductor-Field-Effect-Transistor) consists of two back-to-back p-n junctions, and, transverse to this, a gate-bulk-oxide capacitor. The situation is depicted in Fig. 1.6. If there is no gate voltage ($V_{\rm g} = 0$), then current will not flow at any bias voltage V because necessarily one of the p-n junction will be reverse-biased. The situation changes drastically if the gate is held at a high positive potential $V_{\rm g}$, for then an n-type accumulation layer forms at the bulk-gate interface, thereby connecting source and drain directly and resulting in a gate-controlled current flow. Although not shown in the figure, generally both source and drain are biased positively with respect to the bulk in order to avoid current leakage.

1.2.2 Heterojunctions

Potential uses of a junction formed from two distinct semiconductors were envisioned as early as 1951 by Shockley. Such devices, known as *heterojunctions*, have revolutionized the electronics industry and experimental solid state physics, the latter due to the advent of epitaxial technology which permits growth patterning to nearly atomic precision. Whereas the best inversion layer mobilities in Si MOSFETs are $\mu \approx 2 \times 10^4\,{\rm cm}^2/{\rm V\,s}$, values as high as $10^7\,{\rm cm}^2/{\rm V\,s}$ are pos-

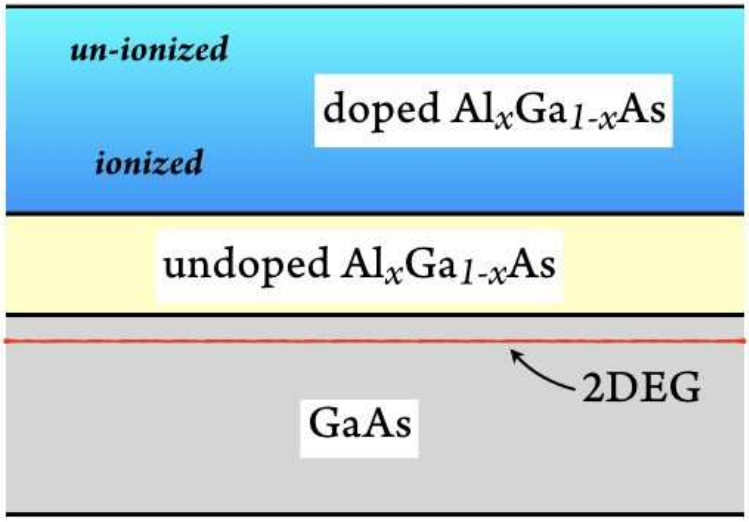


Figure 1.7: GaAs$-$Al$_x$Ga$_{1-x}$As heterojunction.

sible in MBE-fabricated GaAs$-$Al$_x$Ga$_{1-x}$As heterostructures. There are three reasons for this:

(i) MBE (molecular beam epitaxy), as mentioned above, can produce layers which are smooth on an atomic scale. This permits exquisite control of layer thicknesses and doping profiles.

(ii) Use of ternary compounds such as Al$_x$Ga$_{1-x}$As makes for an excellent match in lattice constant across the heterojunction interface, *i.e.* on the order of or better than 1%. By contrast, the Si$-$SiO$_2$ interface is very poor, since SiO$_2$ is a glass.

(iii) By doping the Al$_x$Ga$_{1-x}$As layer far from the interface, Coulomb scattering between inversion layer electrons and dopant ions is suppressed.

Let's consider the chemical potential alignment problem in the case of an n-n heterojunction,

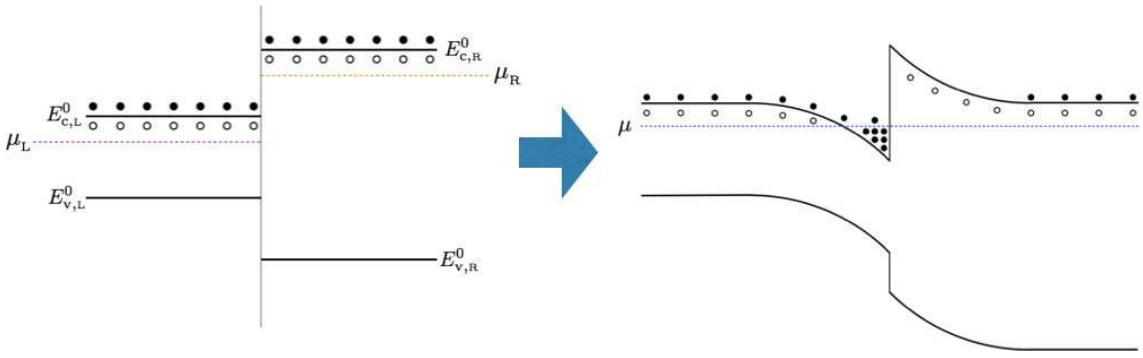


Figure 1.8: Accumulation layer formation in an n-n heterojunction.

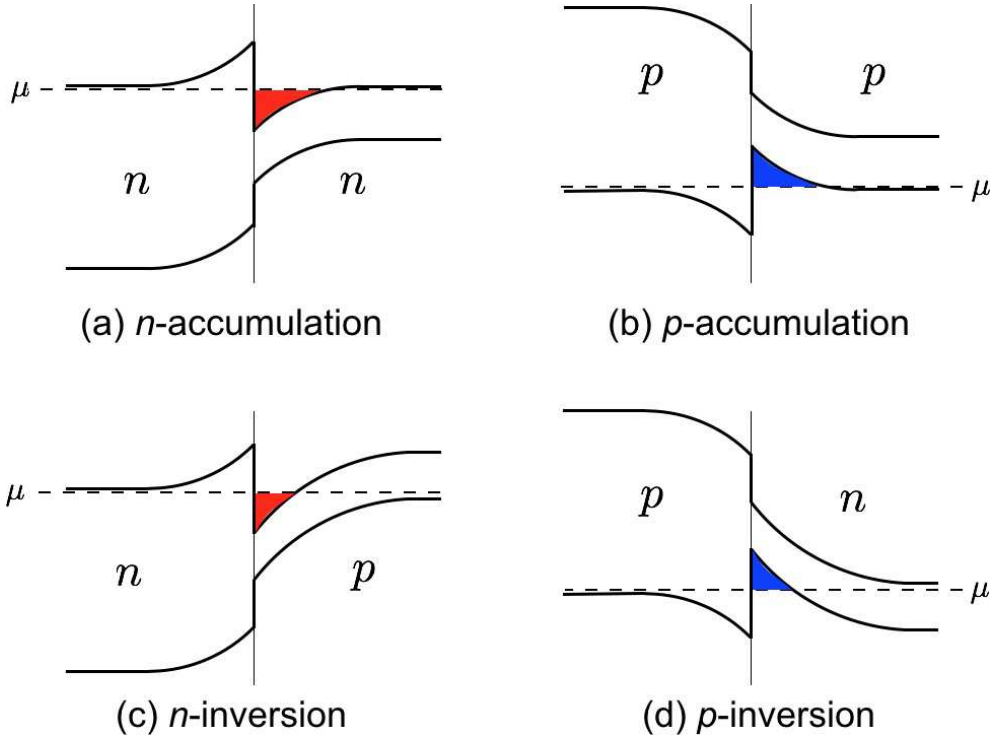


Figure 1.9: Accumulation and inversion in semiconductor heterojunctions. Red regions represent presence of conduction electrons. Blue regions represent presence of valence holes.

sketched in Fig. 1.8. In the GaAs$-$Al$_x$Ga$_{1-x}$As heterojunction, GaAs has the smaller of the two band gaps. Initially there is a mismatch, as depicted in the left panel of the figure. By forming a depletion layer on the side with the larger band gap (Al$_x$Ga$_{1-x}$As), and an accumulation layer on the side with the smaller gap (GaAs), an internal potential $\phi(x)$ is established which aligns the chemical potentials.

Fig. 1.9 shows the phenomena of accumulation and inversion in different possible heterojunctions. There are four possibilities: (a) n-n, (b) p-p, (c) n-p with the n-type material having the larger gap, and (d) n-p with the p-type material having the larger gap.

1.2.3 QM of electron motion normal to 2DEG planes

Consider the case of an n-accumulation or n-inversion layer as depicted in Fig. 1.9. Let the direction perpendicular to the 2DEG be $\hat{z}$, and let the 2DEG lie on the $z > 0$ side of the interface. Assuming that $\hat{z}$ is a principal axis for the effective mass tensor (with eigenvalue m_z), and the magnetic field is along $\pm\hat{z}$, the single electron Hamiltonian is separable into degrees of freedom in the (x, y) plane and those in the $\hat{z}$ direction, *i.e.* $H = H_\perp + H_\parallel$. The eigenvalues and eigenfunctions for $H_\perp$, which governs the planar degrees of freedom, were discussed in §1.3.

Here we consider $H_\parallel$, which we model as

$$H_\parallel = -\frac{\hbar^2}{2m_z}\frac{\partial^2}{\partial z^2} + V(z) \quad , \tag{1.30}$$

with

$$V(z) = -e\phi(z) \approx \begin{cases} 2\pi e\sigma\epsilon^{-1}z & \text{if } z \geq 0 \\ \infty & \text{if } z < 0 \quad . \end{cases} \tag{1.31}$$

Here σ is the 2D charge density of the space charge layer and ϵ the dielectric constant for $z > 0$. Thus, we have a triangular potential.

Next, define the length scale

$$\lambda \equiv \left(\frac{\epsilon\hbar^2}{4\pi\sigma e m_z}\right)^{1/3} \quad , \tag{1.32}$$

the energy scale $\varepsilon_\parallel \equiv \hbar^2/2m_z\lambda^2$, and the dimensionless length $s \equiv z/\lambda$. Then

$$H_\parallel = \varepsilon_\parallel\left(-\frac{\partial^2}{\partial s^2} + s\right) \tag{1.33}$$

with wavefunctions subject to the boundary condition $\varphi(0) = 0$. The solutions are Airy functions. Recall the Airy differential equation,

$$\mathsf{Ai}''(z) - z\,\mathsf{Ai}(z) = 0 \quad . \tag{1.34}$$

Thus, the eigenfunctions of $H_\parallel$ are given by $\varphi_n(z) = \mathsf{Ai}(z + \zeta_n)$, where $\mathsf{Ai}(\zeta_n) = 0$. The first few zeros of $\mathsf{Ai}(z)$ are given by

$$\zeta_1 = -2.3381 \quad , \quad \zeta_2 = -4.0879 \quad , \quad \zeta_3 = -5.5206 \quad , \quad \zeta_4 = -6.7867 \quad , \quad \zeta_5 = -7.9441 \quad . \tag{1.35}$$

The energy eigenvalue corresponding to the eigenfunction $\varphi_n(z)$ is $\mathcal{E}_n = -\zeta_n\,\varepsilon_\parallel$.

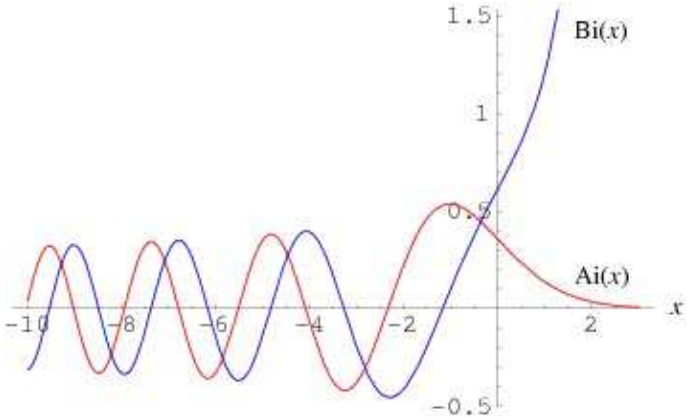


Figure 1.10: Airy functions $\mathsf{Ai}(x)$ and $\mathsf{Bi}(x)$. *Image credit: Wolfram MathWorld.*

1.3 Quantization of Planar Motion

1.3.1 Cyclotron and guiding-center operators

Initially we shall assume spinless (*i.e.* spin-polarized) electrons. Later on we will include effects of the Zeeman term and explore exchange interactions within a Landau level. The single particle Hamiltonian is then

$$H = \frac{1}{2m}\Big(\boldsymbol{p} + \frac{e}{c}\boldsymbol{A}\Big)^2 + V(\boldsymbol{r}) \quad , \tag{1.36}$$

where $V(\boldsymbol{r})$ is the potential. On a toroidal base space, $V(\boldsymbol{r})$ is a doubly periodic function with spatial periods $\boldsymbol{L}_{1,2}$, and $V(\boldsymbol{r}+\boldsymbol{L}_a) = V(\boldsymbol{r})$ for $a = 1, 2$. We assume $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$ is constant[12]. The cyclotron and guiding-center momenta are defined to be

$$\begin{aligned} \boldsymbol{\pi} &= \boldsymbol{p} + \frac{e}{c}\boldsymbol{A} \\ \boldsymbol{\kappa} &= \boldsymbol{p} + \frac{e}{c}\boldsymbol{A} - \frac{e}{c}\boldsymbol{B}\times\boldsymbol{r} \quad . \end{aligned} \tag{1.37}$$

In component notation, we have $\kappa_\mu = p_\mu + \frac{e}{c}A_\mu - \frac{eB}{c}\,\epsilon_{\mu\rho}\,x_\rho$. The commutators are

$$\begin{aligned} \big[\pi_\mu, \pi_\nu\big] &= \frac{e\hbar}{ic}\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) = \frac{i\hbar^2}{\ell^2}\,\epsilon_{\mu\nu} \\ \big[\kappa_\mu, \kappa_\nu\big] &= \frac{e\hbar}{ic}\left(\partial_\mu A_\nu - \partial_\nu A_\mu + 2B\,\epsilon_{\mu\nu}\right) = -\frac{i\hbar^2}{\ell^2}\,\epsilon_{\mu\nu} \\ \big[\pi_\mu, \kappa_\nu\big] &= \frac{e\hbar}{ic}\left(\partial_\mu A_\nu - \partial_\nu A_\mu + B\,\epsilon_{\mu\nu}\right) = 0 \quad , \end{aligned} \tag{1.38}$$

where $\ell = \sqrt{\hbar c/eB}$ is the magnetic length. Now we write

$$\boldsymbol{A} = \tfrac{1}{2}By\,\hat{\boldsymbol{x}} - \tfrac{1}{2}Bx\,\hat{\boldsymbol{y}} - \frac{\hbar c}{e}\,\boldsymbol{\nabla}\chi \quad , \tag{1.39}$$

where $\chi(\boldsymbol{r}) = \chi(\boldsymbol{r}+\boldsymbol{L}_a)$ is an arbitrary gauge function which is periodic on the torus[13].

Now define the complexified operators

$$\begin{aligned} \pi = \pi_x + i\pi_y &= \frac{\hbar}{i}\left(\partial_x + i\partial_y\right) + \frac{eB}{2c}\left(y - ix\right) - \hbar\left(\partial_x + i\partial_y\right)\chi \\ &= \frac{2\hbar}{i}\Big(\bar\partial + \frac{z}{4\ell^2} - i\,\bar\partial\chi\Big) = e^{i\chi}\,e^{-z\bar z/4\ell^2}\,\left(-2i\hbar\,\bar\partial\right)e^{z\bar z/4\ell^2}\,e^{-i\chi} \end{aligned} \tag{1.40}$$

[12] By orienting $\boldsymbol{B}$ along $-\hat{\boldsymbol{z}}$, the non-Gaussian part of the lowest Landau level wavefunctions will be holomorphic in $z = x + iy$, rather than in $\bar z = x - iy$.

[13] To demonstrate the manifest gauge covariance of our description, we shall carry around the gauge function $\chi(\boldsymbol{r})$ for a little while. Students should note on their course evaluations that the professor is sensitive to people with diverse gauge preferences.

and

$$\begin{aligned}\kappa = \kappa_x + i\kappa_y &= \frac{\hbar}{i}\,(\partial_x + i\partial_y) - \frac{eB}{2c}\,(y - ix) - \hbar\,(\partial_x + i\partial_y)\,\chi \\ &= \frac{2\hbar}{i}\Big(\bar\partial - \frac{z}{4\ell^2} - i\,\bar\partial\chi\Big) = e^{i\chi}\,e^{z\bar z/4\ell^2}\,(-2i\hbar\,\bar\partial)\,e^{-z\bar z/4\ell^2}\,e^{-i\chi} \quad ,\end{aligned} \tag{1.41}$$

where $\ell = \sqrt{\hbar c/eB}$ is the magnetic length. We have used $z = x + iy$, $\bar z = x - iy$, in which case

$$\partial = \frac{\partial}{\partial z} = \tfrac{1}{2}\big(\partial_x - i\partial_y\big) \qquad , \qquad \bar\partial = \frac{\partial}{\partial \bar z} = \tfrac{1}{2}\big(\partial_x + i\partial_y\big) \quad . \tag{1.42}$$

Note that

$$\partial^\dagger = -\bar\partial \tag{1.43}$$

under Hermitian conjugation. The commutators of the complexified cyclotron and guiding-center operators are given by

$$\big[\pi, \pi^\dagger\big] = \big[\kappa^\dagger, \kappa\big] = \frac{2\hbar^2}{\ell^2} \quad , \tag{1.44}$$

with $\big[\pi, \kappa\big] = \big[\pi^\dagger, \kappa\big] = 0$. We may now define cyclotron and guiding-center ladder operators,

$$\pi = -\frac{i\sqrt{2}\hbar}{\ell}\,a \qquad , \qquad \kappa = \frac{i\sqrt{2}\hbar}{\ell}\,b^\dagger \tag{1.45}$$

with canonical commutators $[a, a^\dagger] = [b, b^\dagger] = 1$. The kinetic term in the Hamiltonian is then

$$H_0 = \frac{\boldsymbol{\pi}^2}{2m} = \frac{\pi^\dagger\pi}{2m} + \frac{\hbar^2}{2m\ell^2} = \hbar\omega_{\rm c}\big(a^\dagger a + \tfrac{1}{2}\big) \quad , \tag{1.46}$$

Note that H_0 is cyclic in the guiding-center operators, hence each eigenvalue $\varepsilon_n = (n + \frac{1}{2})\hbar\omega_{\rm c}$ is extensively degenerate. As we shall see below, the degeneracy of each of these *Landau levels* is in the thermodynamic limit equal to $N_{\rm L} = BA/\phi_0$, which is the total magnetic flux through the system in units of the Dirac flux quantum.

We may also define the complexified cyclotron and guiding-center coordinates, ξ and $\mathcal{R}$, as follows:

$$\xi = \frac{i\ell^2}{\hbar}\,\pi = \sqrt{2}\ell\,a \qquad , \qquad \mathcal{R} = -\frac{i\ell^2}{\hbar}\,\kappa = \sqrt{2}\ell\,b^\dagger \quad , \tag{1.47}$$

with $[\mathcal{R}, \mathcal{R}^\dagger] = -2\ell^2$ and $[\xi, \xi^\dagger] = 2\ell^2$. Note then that the complexified position $z = x + iy$ is then given by

$$z = \frac{i\ell^2}{\hbar}\,(\pi - \kappa) = \mathcal{R} + \xi = \sqrt{2}\ell\,(a + b^\dagger) \quad , \tag{1.48}$$

with $\bar z = z^\dagger = \sqrt{2}\ell\,(a^\dagger + b)$. For reference, we also have

$$\partial - i\,\partial\chi = \frac{1}{\sqrt{8}\ell}\,(b - a^\dagger) \qquad , \qquad \bar\partial - i\,\bar\partial\chi = \frac{1}{\sqrt{8}\ell}\,(a - b^\dagger) \quad . \tag{1.49}$$

Finally, the following relations may be useful:

$$a = \sqrt{2}\ell\, e^{i\chi}\, e^{-z\bar{z}/4\ell^2}\, \bar{\partial}\, e^{z\bar{z}/4\ell^2}\, e^{-i\chi} \qquad , \qquad b = \sqrt{2}\ell\, e^{i\chi}\, e^{-z\bar{z}/4\ell^2}\, \partial\, e^{z\bar{z}/4\ell^2}\, e^{-i\chi} \tag{1.50}$$

and

$$a^\dagger = -\sqrt{2}\ell\, e^{i\chi}\, e^{z\bar{z}/4\ell^2}\, \partial\, e^{-z\bar{z}/4\ell^2}\, e^{-i\chi} \qquad , \qquad b^\dagger = -\sqrt{2}\ell\, e^{i\chi}\, e^{z\bar{z}/4\ell^2}\, \bar{\partial}\, e^{-z\bar{z}/4\ell^2}\, e^{-i\chi} \quad . \tag{1.51}$$

Exercise : Show that the angular momentum operator satisfies

$$L_z \equiv e^{i\chi}\big(xp_y - yp_x\big)e^{-i\chi} = \hbar\left(b^\dagger b - a^\dagger a\right) \quad . \tag{1.52}$$

1.3.2 Landau level projection

Consider the Hamiltonian $H = H_0 + V(\boldsymbol{r})$ confined to the plane. We may write the potential as a Fourier integral

$$V(\boldsymbol{r}) = \int \frac{d^2k}{(2\pi)^2}\, \hat{V}(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \quad . \tag{1.53}$$

Write $\boldsymbol{r} = \boldsymbol{\mathcal{R}} + \boldsymbol{\xi}$ as a sum over guiding-center and cyclotron coordinates. Since $[\mathcal{R}_\alpha, \xi_\beta] = 0$, we have that

$$\begin{aligned} e^{i\boldsymbol{k}\cdot\boldsymbol{r}} &= e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{\xi}} = e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\, e^{-\boldsymbol{k}^2\ell^2/4}\, e^{i\ell k a^\dagger/\sqrt{2}}\, e^{i\ell\bar{k}a/\sqrt{2}} \\ &= e^{-\boldsymbol{k}^2\ell^2/2}\, e^{i\ell\bar{k}b^\dagger/\sqrt{2}}\, e^{i\ell k b/\sqrt{2}}\, e^{i\ell k a^\dagger/\sqrt{2}}\, e^{i\ell\bar{k}a/\sqrt{2}} \quad . \end{aligned} \tag{1.54}$$

We have skipped a few steps. First, we have written $\boldsymbol{k}\cdot\boldsymbol{\xi} = \mathrm{Re}\,(k\xi^\dagger) = \frac{1}{2}(k\xi^\dagger + \bar{k}\xi)$. Then we wrote $\xi = \sqrt{2}\ell\, a$ and $\xi^\dagger = \sqrt{2}\ell\, a^\dagger$ in terms of the cyclotron latter operators. Finally, we wrote $e^{i\ell(ka^\dagger + \bar{k}a)/\sqrt{2}} = e^{-\boldsymbol{k}^2\ell^2/4}\, e^{i\ell k a^\dagger}\, e^{i\ell\bar{k}a/\sqrt{2}}$ using the Baker-Campbell-Hausdorff equality,

$$e^{A+B} = e^A\, e^B\, e^{-\frac{1}{2}[A,B]} \quad , \tag{1.55}$$

which is true when both A and B commute with their commutator $[A, B]$.

Now suppose we *project* the Hamiltonian onto the $n^{\rm th}$ Landau level. This means we evaluate it's expectation value in the cyclotron oscillator state $|\, n\,\rangle$. The result $\langle\, n\,|\, H\,|\, n\,\rangle$ is still an operator, but only in the space of guiding-center states. In other words, it will only involve the operators b and $b^\dagger$ (or $\mathcal{R}$ and $\mathcal{R}^\dagger$). Now we have to roll up our sleeves and do some work. We have

$$\exp\big(i\ell\bar{k}a/\sqrt{2}\big)\,|\, n\,\rangle = \sum_{j=0}^{n} \frac{1}{j!}\left(\frac{i\ell\bar{k}}{\sqrt{2}}\right)^j a^j\,|\, n\,\rangle = \sum_{j=0}^{n} \frac{1}{j!}\left(\frac{i\ell\bar{k}}{\sqrt{2}}\right)^j \sqrt{\frac{n!}{(n-j)!}}\,|\, n-j\,\rangle \tag{1.56}$$

and so

$$\langle\, n\,|\, \exp\big(i\ell k a^\dagger/\sqrt{2}\big)\exp\big(i\ell\bar{k}a/\sqrt{2}\big)\,|\, n\,\rangle = \sum_{j=0}^{n} \frac{1}{j!}\binom{n}{j}\big(-\tfrac{1}{2}\,\boldsymbol{k}^2\ell^2\big)^j \equiv C_n(\boldsymbol{k}\ell) \quad . \tag{1.57}$$

Therefore,

$$V_n(\boldsymbol{\mathcal{R}}) \equiv \langle\, n \,|\, V \,|\, n \,\rangle = \int \frac{d^2k}{(2\pi)^2}\, \hat{V}(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\, e^{-\boldsymbol{k}^2\ell^2/4}\, C_n(\boldsymbol{k}\ell) \quad . \tag{1.58}$$

Let's examine what happens for the first few values of n:

$$\begin{aligned} C_0(\boldsymbol{k}\ell) &= 1 \\ C_1(\boldsymbol{k}\ell) &= 1 - \tfrac{1}{2}(\boldsymbol{k}\ell)^2 \\ C_2(\boldsymbol{k}\ell) &= 1 - (\boldsymbol{k}\ell)^2 + \tfrac{1}{8}\,(\boldsymbol{k}\ell)^4 \quad , \end{aligned} \tag{1.59}$$

where by $(\boldsymbol{k}\ell)^4$ we mean $|\boldsymbol{k}|^4\ell^4$. Multiplying by the $\exp(-\frac{1}{4}\boldsymbol{k}^2\ell^2)$ factor, and expanding in powers of $\boldsymbol{k}$, we see that the projected potential $V_n(\boldsymbol{\mathcal{R}})$ is given by $V(\boldsymbol{\mathcal{R}})$ plus a series of corrections which can be expressed in terms of powers of the Laplacian ∇^2 acting on $V(\boldsymbol{\mathcal{R}})$. For example,

$$V_0(\boldsymbol{\mathcal{R}}) = V(\boldsymbol{\mathcal{R}}) + \tfrac{1}{4}\,\ell^2\,\nabla^2 V(\boldsymbol{\mathcal{R}}) + \ldots \quad , \tag{1.60}$$

which is the quantum analog of Eqn. 1.28.

Some words of caution are appropriate here. Since $[\mathcal{R}_\alpha, \mathcal{R}_\beta] = -i\ell^2\epsilon_{\alpha\beta}$, we must not be cavalier regarding operator order. To be safe, we might choose to express $V_n(\boldsymbol{\mathcal{R}})$ in some canonical form, such as the *normal ordered* form, in which all the $\mathcal{R}^\dagger$ operators appear to the right of all $\mathcal{R}$ operators[14]. That is, we write

$$\begin{aligned} \exp(i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}) &= \exp(\tfrac{i}{2}\bar{k}\mathcal{R})\, \exp(\tfrac{i}{2}k\mathcal{R}^\dagger)\, \exp\big(\tfrac{1}{8}\boldsymbol{k}^2[\mathcal{R},\mathcal{R}^\dagger]\big) \\ &= \exp(-\tfrac{1}{4}\boldsymbol{k}^2\ell^2)\, \exp(\tfrac{i}{2}\bar{k}\mathcal{R})\, \exp(\tfrac{i}{2}k\mathcal{R}^\dagger) \end{aligned} \tag{1.61}$$

in the integrand of Eqn. 1.58. Also, it goes without saying that $\langle\, n \,|\, H_0 \,|\, n \,\rangle = (n + \frac{1}{2})\hbar\omega_{\rm c}$. But I suppose I said it anyway.

1.3.3 Landau level mixing

It is apparent that the Hamiltonian $H = H_0 + V(\boldsymbol{r})$ may be written as

$$\begin{aligned} H &= \sum_{n,n=0}^{\infty} H_{nn'}(\boldsymbol{\mathcal{R}})\,|\, n \,\rangle\langle\, n' \,| \\ H_{nn'}(\boldsymbol{\mathcal{R}}) &= (n + \tfrac{1}{2})\hbar\omega_{\rm c}\,\delta_{nn'} + V_{nn'}(\boldsymbol{\mathcal{R}}) \quad , \end{aligned} \tag{1.62}$$

where, for each $\{n, n'\}$, $V_{nn'}(\boldsymbol{\mathcal{R}}) \equiv \langle\, n \,|\, V \,|\, n' \,\rangle$ is an operator in the space of guiding-center degrees of freedom, given by

$$V_{nn'}(\boldsymbol{\mathcal{R}}) = \int \frac{d^2k}{(2\pi)^2}\, \hat{V}(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\, e^{-\boldsymbol{k}^2\ell^2/4}\, \langle\, n \,|\, \exp\big(i\ell k a^\dagger/\sqrt{2}\big)\, \exp\big(i\ell\bar{k}a/\sqrt{2}\big)\,|\, n' \,\rangle \quad . \tag{1.63}$$

[14] And, hence, all guiding-center annihilation operators b appear to the right of all guiding-center creation operators $b^\dagger$.

The off-diagonal terms describe *Landau level mixing* processes. For example, if we retain only the $n = 0$ and $n = 1$ LLs, we have the truncated Hamiltonian

$$H_{\rm trunc} = \begin{pmatrix} \frac{1}{2}\hbar\omega_{\rm c} & 0 \\ 0 & \frac{3}{2}\hbar\omega_{\rm c} \end{pmatrix} + \int \frac{d^2k}{(2\pi)^2}\,\hat{V}(\boldsymbol{k})\,e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\,e^{-\boldsymbol{k}^2\ell^2/4}\begin{pmatrix} 1 & -i\ell\bar{k}/\sqrt{2} \\ i\ell k/\sqrt{2} & 1-\frac{1}{2}\boldsymbol{k}^2\ell^2 \end{pmatrix} \quad . \tag{1.64}$$

Since there is a gap of $\hbar\omega_{\rm c}$ between consecutive Landau levels, LL mixing is usually treated perturbatively.

1.3.4 The lowest Landau level

The eigenvalue of $a^\dagger a$ is an integer which corresponds to the Landau level index. For states in the lowest Landau level, we have

$$a\,\psi(\boldsymbol{r}) = 0 \quad \Longrightarrow \quad \psi(\boldsymbol{r}) = e^{i\chi(\boldsymbol{r})}\,e^{-\boldsymbol{r}^2/4\ell^2}\,f(z) \quad , \tag{1.65}$$

where $z = x + iy$. At this point, $f(z)$ is any analytic function. As we shall soon see, periodicity on the torus further constrains the form of $f(z)$.

In zero magnetic field, the density of states (per unit area, per unit energy) is constant:

$$g(\varepsilon, B=0)\,d\varepsilon = \frac{d^2k}{(2\pi)^2} = \frac{k\,dk}{2\pi} \quad \Rightarrow \quad g(\varepsilon) = \frac{m}{2\pi\hbar^2} \tag{1.66}$$

since $\varepsilon = \hbar^2k^2/2m$. When B is finite, the spectrum collapses into discrete Landau levels with energies $\varepsilon_n = (n+\frac{1}{2})\hbar\omega_{\rm c}$. The density of states is

$$g(\varepsilon, B) = \frac{B}{\phi_0}\sum_{n=0}^{\infty}\delta\big(\varepsilon - (n+\tfrac{1}{2})\hbar\omega_{\rm c}\big) \quad . \tag{1.67}$$

The number of Landau levels below energy E is $E/\hbar\omega_{\rm c}$, rounded to the nearest integer. To check the coefficient B/ϕ_0 in the above expression note that the total number of states per unit area below energy E is then

$$\frac{B}{\phi_0}\cdot\frac{E}{\hbar\omega_{\rm c}} = \frac{mE}{2\pi\hbar^2} \quad , \tag{1.68}$$

which agrees with the $B = 0$ result. Below, we shall count the number of states precisely using a toroidal geometry.

We define the wavefunction $\psi_0(\boldsymbol{r})$ to satisfy $a\psi_0 = b\psi_0 = 0$. Imposing normalization,

$$\psi_0(\boldsymbol{r}) = (2\pi\ell^2)^{-1/2}\,e^{i\chi(\boldsymbol{r})}\,e^{-z\bar{z}/4\ell^2} \quad . \tag{1.69}$$

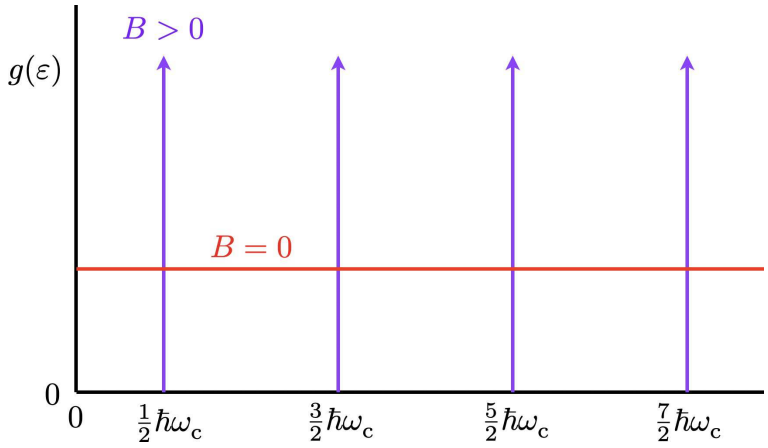


Figure 1.11: Density of states in $d = 2$ for $B = 0$ (red) and $B > 0$ (purple).

A complete and orthonormal set of wavefunctions is given by the collection

$$\begin{aligned}\psi_{m,n}(\boldsymbol{r}) &= \frac{(a^\dagger)^n}{\sqrt{n!}}\,\frac{(b^\dagger)^m}{\sqrt{m!}}\,\psi_0(\boldsymbol{r}) = \frac{\left(-\sqrt{2}\,\ell\right)^{m+n}}{\sqrt{2\pi\ell^2 m!n!}}\,e^{i\chi(\boldsymbol{r})}\,e^{z\bar{z}/4\ell^2}\,\partial^n\,\bar{\partial}^m\,e^{-z\bar{z}/2\ell^2} \\ &= (-1)^n\sqrt{\frac{n!}{2\pi\ell^2 m!}}\left(\frac{z}{\sqrt{2}\ell}\right)^{m-n} L_n^{(m-n)}(z\bar{z}/2\ell^2)\,e^{-z\bar{z}/4\ell^2}e^{i\chi(\boldsymbol{r})} \quad ,\end{aligned} \tag{1.70}$$

where $L_n^{(\alpha)}(x) = \frac{1}{n!}\,x^{-\alpha}\,e^x\,\frac{d^n}{dx^n}\left(x^{n+\alpha}\,e^{-x}\right)$ is an associated Laguerre polynomial. Note that, per Eqn. 1.52, this is also an eigenbasis of angular momentum, *viz.*

$$L_z\,|\,m,n\,\rangle = \hbar(b^\dagger b - a^\dagger a)\,|\,m,n\,\rangle = \hbar(m-n)\,|\,m-n\,\rangle \quad . \tag{1.71}$$

Completeness entails the relation $\sum_{m=0}^\infty\sum_{n=0}^\infty |\,m,n\,\rangle\langle\,m,n\,| = 1$, which is to say

$$\sum_{m=0}^\infty\sum_{n=0}^\infty \psi_{m,n}(\boldsymbol{r})\,\psi^*_{m,n}(\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}') \quad . \tag{1.72}$$

Note, however, that if we sum over only the states in the lowest Landau level, then

$$\begin{aligned}\sum_{m=0}^\infty \psi_{m,0}(\boldsymbol{r})\,\psi^*_{m,0}(\boldsymbol{r}') &= \frac{e^{i\chi(\boldsymbol{r})}e^{-i\chi(\boldsymbol{r}')}}{2\pi\ell^2}\sum_{m=0}^\infty\frac{1}{m!}\left(\frac{z\bar{z}'}{2\ell^2}\right)^m e^{-z\bar{z}/4\ell^2}e^{-z'\bar{z}'/4\ell^2} \\ &= \frac{e^{i\chi(\boldsymbol{r})}e^{-i\chi(\boldsymbol{r}')}}{2\pi\ell^2}\exp\left(-\frac{z\bar{z}}{4\ell^2}-\frac{z'\bar{z}'}{4\ell^2}+\frac{z\bar{z}'}{2\ell^2}\right) \\ &= \frac{e^{i\chi(\boldsymbol{r})}e^{-i\chi(\boldsymbol{r}')}}{2\pi\ell^2}\exp\left(-\frac{|z-z'|^2}{4\ell^2}+i\,\frac{\mathsf{Im}\,z\bar{z}'}{2\ell^2}\right) \\ &= \frac{e^{i\chi(\boldsymbol{r})}e^{-i\chi(\boldsymbol{r}')}}{2\pi\ell^2}\exp\left(-\frac{(\boldsymbol{r}-\boldsymbol{r}')^2}{4\ell^2}-i\,\frac{\boldsymbol{r}\times\boldsymbol{r}'\cdot\hat{\boldsymbol{z}}}{2\ell^2}\right)\end{aligned} \tag{1.73}$$

rather than $\delta(\boldsymbol{r}-\boldsymbol{r}')$. This tells us that the shortest distance scale on which we can localize an electron in the lowest Landau level is the magnetic length ℓ.

In the lowest Landau level (LLL), we may write

$$\psi_m(\boldsymbol{r}) = \frac{e^{i\chi(\boldsymbol{r})}}{\sqrt{2\pi\ell^2\, m!}} \left(\frac{z}{\sqrt{2}\ell}\right)^m e^{-z\bar{z}/4\ell^2} \quad . \tag{1.74}$$

This is generally known as the *angular momentum basis*.

1.3.5 Landau strip basis

Had we instead chosen the gauge $\boldsymbol{A} = -Bx\hat{\boldsymbol{y}}$ (again corresponding to $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$), then

$$H_0 = \frac{p_x^2}{2m} + \frac{(p_y - \frac{eB}{c}\, x)^2}{2m} \quad . \tag{1.75}$$

There is now translational invariance along $\hat{\boldsymbol{y}}$, hence the wavefunctions may be written as $\psi(x,y) = e^{ik_y y}\,\phi(x)$, with $\phi(x)$ an eigenfunction of

$$H_0(k_y) = \frac{p_x^2}{2m} + \frac{(\hbar k_y - \frac{eB}{c}\, x)^2}{2m} = \frac{p_x^2}{2m} + \tfrac{1}{2}m\omega_{\rm c}^2\,(x - \ell^2 k_y)^2 \quad . \tag{1.76}$$

This is the one-dimensional harmonic oscillator, with eigenfunctions $\phi_n(x-\ell^2 k_y)$, where

$$\phi_n(x) = \frac{1}{\sqrt{2^n n!}}\,(\pi\ell^2)^{-1/4}\,\mathsf{H}_n(x/\ell)\,e^{-x^2/2\ell^2} \quad , \tag{1.77}$$

where $\mathsf{H}_n(u)$ is the $n^{\rm th}$ Hermite polynomial, and corresponding eigenvalues $\varepsilon_n = (n+\frac{1}{2})\hbar\omega_{\rm c}$. The full basis set of wavefunctions as a function of (x,y) is labeled by a discrete Landau level index n and a continuous index k_y, *viz.*

$$\psi_{n,k_y}(x,y) = L_y^{-1/2}\,e^{ik_y y}\,\phi_n(x-\ell^2 k_y) \quad . \tag{1.78}$$

On a cylinder $\mathbb{R}\times\mathbb{S}^1$ where $y\in[0,L_y]$, periodic boundary conditions requires k_y to be quantized due to the relation $e^{ik_y L_y}=1$.

1.3.6 Magnetic translation operators

The magnetic translation operators (MTOs) are defined as

$$t(\boldsymbol{d}) = \exp(i\boldsymbol{\kappa}\cdot\boldsymbol{d}/\hbar) = \exp\big[(db - \bar{d}b^\dagger)/\sqrt{2}\ell\big] \quad . \tag{1.79}$$

For each $\boldsymbol{d}$, $t^{-1}(\boldsymbol{d}) = t^\dagger(\boldsymbol{d})$, *i.e.* each MTO is a unitary operator. Note also that $\big[t(\boldsymbol{d}), \boldsymbol{\pi}\big] = 0$, so the magnetic translations commute with the kinetic energy, $H_0 = \frac{\pi^2}{2m}$. Acting on any function of the coordinates, we have

$$t(\boldsymbol{d})\,\psi(\boldsymbol{r})\,t^\dagger(\boldsymbol{d}) = \psi(\boldsymbol{r}+\boldsymbol{d}) \quad , \tag{1.80}$$

which is why $t(\boldsymbol{d})$ is a translation operator. It is a worthwhile exercise for the student to show that while $t_0(\boldsymbol{d})\,\psi(\boldsymbol{r}) = \psi(\boldsymbol{r}+\boldsymbol{d})$, where $t_0(\boldsymbol{d}) = \exp(i\boldsymbol{p}\cdot\boldsymbol{d}/\hbar) = \exp(\boldsymbol{d}\cdot\boldsymbol{\nabla})$ is the translation operator with $\boldsymbol{B} = 0$, for the MTOs we have

$$t(\boldsymbol{d})\,\psi(\boldsymbol{r}) = e^{i\chi(\boldsymbol{r})}\,e^{-i\chi(\boldsymbol{r}+\boldsymbol{d})}\,e^{-i\boldsymbol{d}\times\boldsymbol{r}\cdot\hat{\boldsymbol{z}}/2\ell^2}\,\psi(\boldsymbol{r}+\boldsymbol{d}) \quad . \tag{1.81}$$

Due to the magnetic field, two arbitrary magnetic translations do not necessarily commute. Rather,

$$t(\boldsymbol{d}_1)\,t(\boldsymbol{d}_2) = e^{i\hat{\boldsymbol{z}}\cdot\boldsymbol{d}_1\times\boldsymbol{d}_2/2\ell^2}\,t(\boldsymbol{d}_1+\boldsymbol{d}_2) = e^{i\hat{\boldsymbol{z}}\cdot\boldsymbol{d}_1\times\boldsymbol{d}_2/\ell^2}\,t(\boldsymbol{d}_2)\,t(\boldsymbol{d}_1) \quad . \tag{1.82}$$

Thus, $\big[t(\boldsymbol{d}_1), t(\boldsymbol{d}_2)\big] = 0$ if and only if $\hat{\boldsymbol{z}}\cdot\boldsymbol{d}_1\times\boldsymbol{d}_2 = 2\pi\ell^2 q$, where q is an integer.

1.3.7 Coherent state wavefunctions

Having tired of carrying the stupid gauge function $\chi(\boldsymbol{r})$ with us for so long, we will now drop it[15], which means working in the symmetric gauge, with $\chi(\boldsymbol{r}) = 0$. Consider again the MTO,

$$\begin{aligned} t(\boldsymbol{R}) &= \exp(i\boldsymbol{\kappa}\cdot\boldsymbol{R}/\hbar) = e^{(Rb-\bar{R}b^\dagger)/\sqrt{2}\ell} \\ &= e^{-R\bar{R}/4\ell^2}e^{-\bar{R}b^\dagger/\sqrt{2}\ell}\,e^{-Rb/\sqrt{2}\ell} \quad , \end{aligned} \tag{1.83}$$

where we have again invoked BCH. The LLL coherent state wavefunction is defined to be[16]

$$|\,\boldsymbol{R}\,\rangle = t^\dagger(\boldsymbol{R})|\,0\,\rangle = e^{-R\bar{R}/4\ell^2}e^{\bar{R}b^\dagger/\sqrt{2}\ell}\,|\,0,0\,\rangle \tag{1.84}$$

It is left as an exercise to the reader to verify the following formulae:

$$\begin{aligned} \varphi_{\boldsymbol{R}}(\boldsymbol{r}) = \langle\,\boldsymbol{r}\,|\,t^\dagger(\boldsymbol{R})\,|\,0\,\rangle &= \frac{1}{\sqrt{2\pi\ell^2}}\,e^{-i\boldsymbol{r}\times\boldsymbol{R}\cdot\hat{\boldsymbol{z}}/2\ell^2}e^{-(\boldsymbol{r}-\boldsymbol{R})^2/4\ell^2} \\ \langle\,\boldsymbol{R}\,|\,\boldsymbol{R}'\,\rangle &= \exp\left(-\frac{(\boldsymbol{R}-\boldsymbol{R}')^2}{4\ell^2} - i\,\frac{\boldsymbol{R}\times\boldsymbol{R}'\cdot\hat{\boldsymbol{z}}}{2\ell^2}\right) \\ \int\frac{d^2R}{2\pi\ell^2}\,|\,\boldsymbol{R}\,\rangle\langle\,\boldsymbol{R}\,| &= \sum_{m=0}^{\infty}|\,m,0\,\rangle\langle\,m,0\,| \equiv \Pi_0 \quad . \end{aligned} \tag{1.85}$$

[15] As we are feeling persons, we hope that the gauge function manages to find its way back home to its loved ones, whom it can regale with heroic stories of having served alongside other important factors in various operators and wavefunctions.

[16] When confining our attention to the LLL, it is convenient to drop the cyclotron quantum number and write the guiding center vacuum state simply as $|\,0\,\rangle$ rather than $|\,0,0\,\rangle$.

Thus, the coherent state wavefunction $\varphi_{\boldsymbol{R}}(\boldsymbol{r})$ is Gaussianly localized about $\boldsymbol{r} = \boldsymbol{R}$, but contains phase information as well. The coherent states admit a resolution of unity within the lowest Landau level, or indeed for any Landau level if we define $|\,\boldsymbol{R}, n\,\rangle \equiv t^\dagger(\boldsymbol{R})\,|\,0, n\,\rangle$, in which case

$$\int \frac{d^2R}{2\pi\ell^2}\,|\,\boldsymbol{R}, n\,\rangle\langle\,\boldsymbol{R}, n\,| = \sum_{m=0}^{\infty} |\,m, n\,\rangle\langle\,m, n\,| \equiv \Pi_n \tag{1.86}$$

is the projector onto the $n^{\rm th}$ Landau level.

1.4 Landau Levels in Graphene

1.4.1 Quick overview

First we will skip all the details and quickly derive the spectrum of the Landau levels in graphene, which is different than for the case in ballistic 2DEGs. Then we will circle back and build a model of what graphene actually *is* at the atomic level and validate everything we find. We'll even allow for lattice *strain* and see how it can generate a *pseudomagnetic field* which in experiments can be on the order of 300 Tesla!

The two-dimensional Dirac Hamiltonian is

$$H = \hbar v_{\rm F}\,\boldsymbol{\sigma}\cdot\boldsymbol{q} = \hbar v_{\rm F}\begin{pmatrix} 0 & q_x - iq_y \\ q_x + iq_y & 0 \end{pmatrix} \quad , \tag{1.87}$$

where $v_{\rm F} \approx c/300$ is the Fermi velocity in graphene. At this point, the electron spin is a spectator, and the Pauli matrix structure here is in the *isospin* space associated with the two triangular sublattices of the hexagonal graphene structure. To describe the Zeeman term, or, more interestingly, spin-orbit effects, we would need to invoke a second set of Pauli matrices (τ^x, τ^y, τ^z) acting on the spin degrees of freedom[17]. In the presence of a magnetic field, the minimal coupling prescription applies[18], and we have

$$H(\boldsymbol{A}) = v_{\rm F}\begin{pmatrix} 0 & \pi^\dagger \\ \pi & 0 \end{pmatrix} \tag{1.88}$$

where $\pi = \hbar(q_x + iq_y) + \frac{e}{c}(A_x + iA_y)$ as before. We then have

$$[\pi, \pi^\dagger] = \frac{2\hbar e}{c}\big(\partial_y A_x - \partial_x A_y\big) = -\frac{\hbar e B_z}{c} \quad , \tag{1.89}$$

[17] In graphene, spin-orbit effects are very weak due to the low Z value of carbon. In higher Z systems, spin-orbit effects are crucial and may give rise to topological insulator behavior.

[18] The validity of the Peierls substitution $E_n(\boldsymbol{k}) \to E_n(\boldsymbol{k} + \frac{e}{\hbar c}\boldsymbol{A})$ for Bloch electrons is a nontrivial matter and was first established by W. Kohn, *Phys. Rev.* **115**, 1460 (1959).

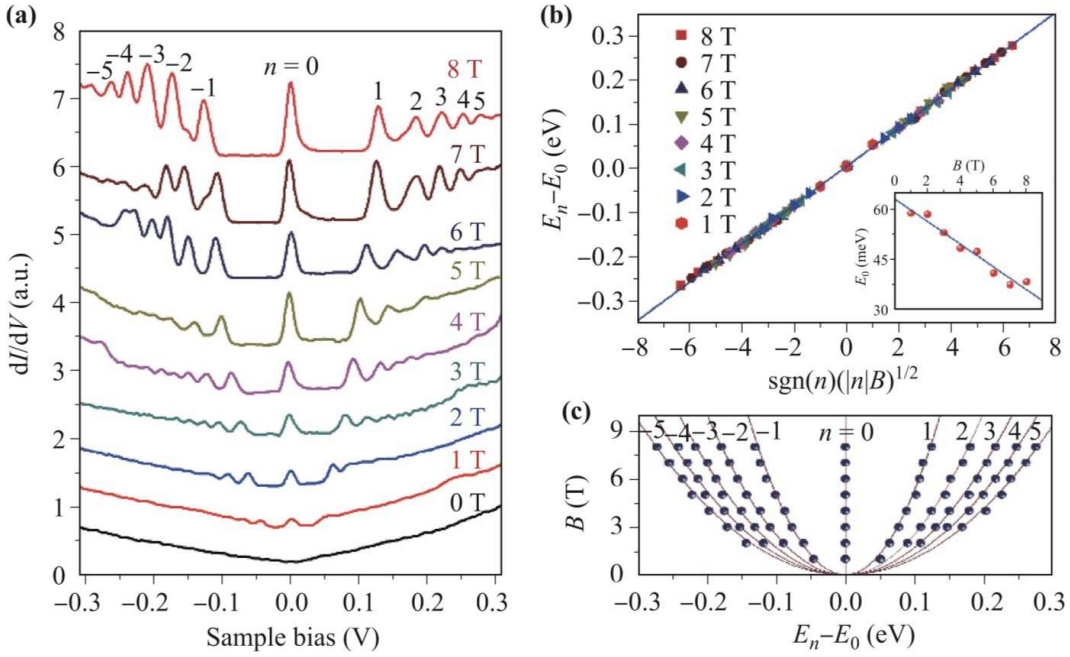


Figure 1.12: Landau levels in monolayer graphene, as reported by L.-J. Yin *et al.*, *Front. Phys.* **12**, 127208 (2017). Left: Differential conductance spectra of a graphene monolayer on a graphite surface as a function of bias voltage and for several values of magnetic field. Right: The LL dispersion is in excellent agreement with the 2D Dirac theory.

and assuming $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$, we again recover $[\pi, \pi^\dagger] = 2\hbar^2\ell^{-2}$ with $\ell = \sqrt{\hbar c/eB}$ the magnetic length. Defining the cyclotron ladder operators $a = \ell\pi/\sqrt{2}\,\hbar$ and $a^\dagger = \ell\pi^\dagger/\sqrt{2}\,\hbar$ as before[19], we have

$$H = \frac{\sqrt{2}\hbar v_{\rm F}}{\ell}\begin{pmatrix} 0 & a^\dagger \\ a & 0 \end{pmatrix} \quad . \tag{1.90}$$

Note that the guiding center operators b and $b^\dagger$ are cyclic in H, hence there is an extensive degeneracy of each Landau level.

It is easy to see that the eigenvectors of H, expressed in terms of the cyclotron oscillator states $|\,n\,\rangle$, are given by

$$|\,\Psi_0\,\rangle = \begin{pmatrix} |\,0\,\rangle \\ 0 \end{pmatrix} \qquad , \qquad |\,\Psi_{\pm,n}\,\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} |\,n\,\rangle \\ \pm|\,n-1\,\rangle \end{pmatrix} \tag{1.91}$$

where $n \in \mathbb{Z}_+ = \{1, 2, 3, \ldots\}$. The corresponding eigenvalues are

$$E_0 = 0 \qquad , \qquad E_{\pm,n} = \pm\sqrt{2n}\,\frac{\hbar v_{\rm F}}{\ell} \quad . \tag{1.92}$$

[19]Not quite as before. In Eqn. 1.45 there is an extra factor of i which we find convenient to remove here.

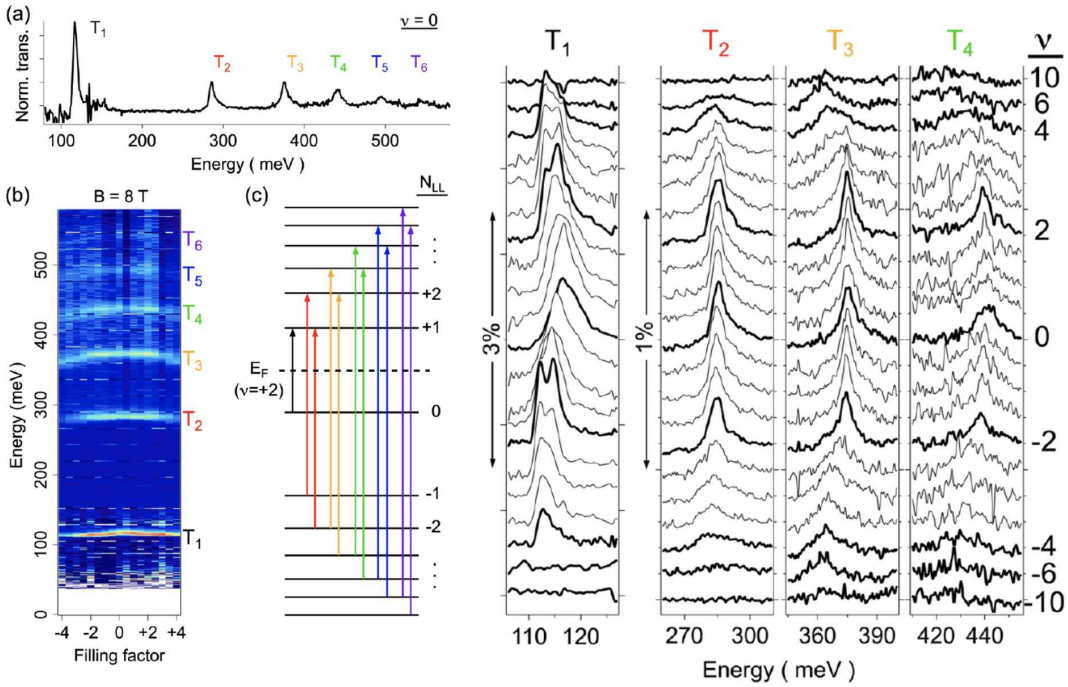


Figure 1.13: Infrared cyclotron resonance data on high mobility monolayer graphene by B. J. Russell *et al.*, *Phys. Rev. Lett.* **120**, 047401 (2018). The interband LL transitions are measured at fixed $B = 8T$. The density n and thus the filling fraction ν are varied by an applied gate voltage. Left: (a) Normalized transition at $\nu = 0$, showing the first six resonances. (b) Color map of resonances *vs.* filling fraction, showing small variations even at fixed ω_c . (c) Schematic of inter-LL transitions. Right: Dependence of the first four inter-LL transitions on filling fraction. Note each filled LL index corresponds to $\Delta\nu = 2$ because of spin degeneracy.

We now note several important distinctions between this spectrum and that for the ballistic case derived in §1.3:

(i) The LL spectrum is particle-hole symmetric, with a zero energy LL and symmetrically arranged levels at $E_{-,n} = -E_{+,n}$.

(ii) Rather than the ballistic LL spectrum $E_n = (n + \frac{1}{2})\hbar\omega_c$ where $\omega_c = eB/mc$, the Dirac LL spectrum is not evenly spaced nor is it linear in B. Instead, energies grow with LL index and field as $\sqrt{|nB|}$.

It turns out that this is only half the story, because there is another Dirac point in the graphene band structure where the corresponding long wavelength Hamiltonian is $H' = \sigma^x H \sigma^x$. The eigenspectra of these two 2D Dirac Hamiltonians are therefore identical.

Scanning probe measurements in monolayer graphene shown in Fig. **??** reveal an excellent fit to the Dirac dispersion of Eqn. 1.92. More careful investigations, however, from infrared

cyclotron resonance (Fig. 1.13) show small variations in inter-LL resonance energies at fixed B as a function of the filling fraction ν. These are attributed to interaction effects, which we shall study later on in these notes.

OK, now let's derive all this stuff from scratch.

1.4.2 Direct and reciprocal lattice

Graphene is a two-dimensional form of pure carbon arranged in a honeycomb lattice, where each site is threefold coordinated. The electronic configuration of C is $[\mathrm{He}]\,2\mathrm{s}^2\,2\mathrm{p}^2$. The $2\mathrm{s}$ and $2\mathrm{p}$ orbitals engage in sp^2 hybridization, forming on each carbon atom three planar orbitals oriented 120° from each other. These engage in covalent bonding with each of a given C atom's three neighbors. The remaining fourth electron is in a p_z state – the so-called π orbital. The simplest model of graphene considers as inert the [He] core and the covalently bonded sp^2 orbitals and focuses on the remaining single π electron per site.

The honeycomb lattice is a triangular Bravais lattice with a two element basis. The Bravais lattice sites are located at $\boldsymbol{R} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2$, with elementary direct lattice vectors

$$\boldsymbol{a}_1 = a\left(\tfrac{1}{2}\hat{\boldsymbol{x}} - \tfrac{\sqrt{3}}{2}\hat{\boldsymbol{y}}\right) \quad , \quad \boldsymbol{a}_2 = a\left(\tfrac{1}{2}\hat{\boldsymbol{x}} + \tfrac{\sqrt{3}}{2}\hat{\boldsymbol{y}}\right) \quad , \tag{1.93}$$

as shown in the left panel of Fig. 1.14. The basis vectors are then 0 and $\boldsymbol{\delta}_1 = \delta\hat{\boldsymbol{y}}$, where $\delta = a/\sqrt{3}$ is the spacing between C atoms. It is useful to define the two other nearest neighbor vectors $\boldsymbol{\delta}_{2,3}$ as shown in the figure, in which case

$$\boldsymbol{\delta}_1 = \tfrac{1}{3}(-\boldsymbol{a}_1 + \boldsymbol{a}_2) \quad , \quad \boldsymbol{\delta}_2 = \tfrac{1}{3}(2\boldsymbol{a}_1 + \boldsymbol{a}_2) \quad , \quad \boldsymbol{\delta}_3 = \tfrac{1}{3}(-\boldsymbol{a}_1 - 2\boldsymbol{a}_2) \quad . \tag{1.94}$$

Note that $\boldsymbol{\delta}_1 + \boldsymbol{\delta}_2 + \boldsymbol{\delta}_3 = 0$. The reciprocal lattice is triangular, with elementary reciprocal lattice vectors

$$\boldsymbol{b}_1 = \frac{4\pi}{a\sqrt{3}}\left(\tfrac{\sqrt{3}}{2}\hat{\boldsymbol{x}} - \tfrac{1}{2}\hat{\boldsymbol{y}}\right) \quad , \quad \boldsymbol{b}_2 = \frac{4\pi}{a\sqrt{3}}\left(\tfrac{\sqrt{3}}{2}\hat{\boldsymbol{x}} + \tfrac{1}{2}\hat{\boldsymbol{y}}\right) \quad . \tag{1.95}$$

Note that $\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 2\pi\delta_{ij}$. Let the tight binding hopping matrix elements along the $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, and $\boldsymbol{\delta}_3$ directed links be $-t_1$, $-t_2$, and $-t_3$, respectively. Writing the fermionic creation operators for electrons at the A and B sites in unit cell $\boldsymbol{R}$ as $a^\dagger_{\boldsymbol{R}}$ and $b^\dagger_{\boldsymbol{R}}$, the tight binding Hamiltonian is then

$$\begin{aligned} H &= -\sum_{\boldsymbol{R}} \Big\{ \left(t_1\, a^\dagger_{\boldsymbol{R}}\, b_{\boldsymbol{R}} + t_2\, a^\dagger_{\boldsymbol{R}}\, b_{\boldsymbol{R}+\boldsymbol{a}_1} + t_3\, a^\dagger_{\boldsymbol{R}}\, b_{\boldsymbol{R}-\boldsymbol{a}_2}\right) + \mathrm{H.c.} \Big\} \\ &= -t\sum_{\boldsymbol{k}} \Big\{ a^\dagger_{\boldsymbol{k}}\, b_{\boldsymbol{k}} \left(u + v\, e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + w\, e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right) + \mathrm{H.c.} \Big\} \quad , \end{aligned} \tag{1.96}$$

where we have written $t_1 \equiv ut$, $t_2 \equiv vt$, and $t_3 \equiv wt$, with t the fundamental hopping energy scale and u, v, and w all dimensionless. We have diagonalized H within each crystal momentum $\boldsymbol{k}$ sector via the relations

$$a_{\boldsymbol{R}} = \frac{1}{\sqrt{N_{\mathrm{c}}}}\sum_{\boldsymbol{k}} a_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \quad , \quad a_{\boldsymbol{k}} = \frac{1}{\sqrt{N_{\mathrm{c}}}}\sum_{\boldsymbol{R}} a_{\boldsymbol{R}}\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \quad , \tag{1.97}$$

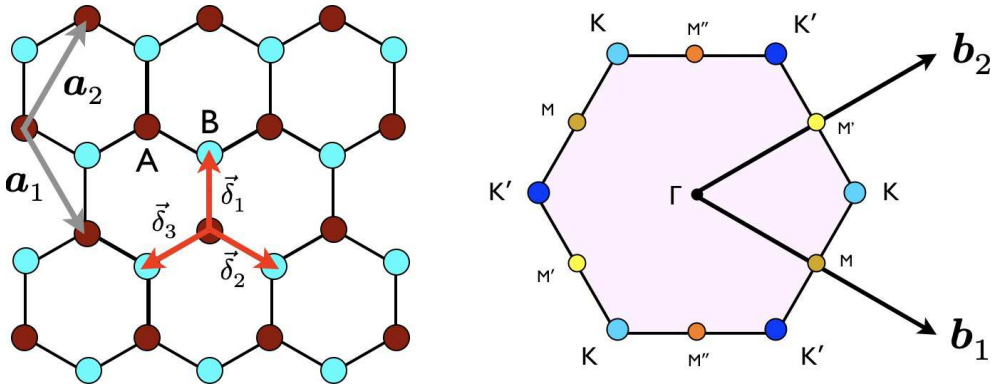


Figure 1.14: The honeycomb lattice is a triangular Bravais lattice with a two element basis (A and B). Left: Real space lattice, with elementary direct lattice vectors $\boldsymbol{a}_{1,2}$. Right: Brillouin zone, with elementary reciprocal lattice vectors $\boldsymbol{b}_{1,2}$ and high symmetry points identified.

where $N_{\rm c}$ is the number of unit cells, and with corresponding definitions on the B sites. Thus $\{a_{\boldsymbol{R}}, a^\dagger_{\boldsymbol{R}'}\} = \delta_{\boldsymbol{R},\boldsymbol{R}'}$ and $\{a_{\boldsymbol{k}}, a^\dagger_{\boldsymbol{k}'}\} = \delta_{\boldsymbol{k},\boldsymbol{k}'}$. We now define

$$\gamma_{\boldsymbol{k}} \equiv u + v\, e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + w\, e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2} \quad , \tag{1.98}$$

which allows us to write

$$H = -t \sum_{\boldsymbol{k}} \begin{pmatrix} a^\dagger_{\boldsymbol{k}} & b^\dagger_{\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} 0 & \gamma_{\boldsymbol{k}} \\ \gamma^*_{\boldsymbol{k}} & 0 \end{pmatrix} \begin{pmatrix} a_{\boldsymbol{k}} \\ b_{\boldsymbol{k}} \end{pmatrix} \quad . \tag{1.99}$$

The energy eigenvalues are this $E_{\boldsymbol{k},\pm} = \pm t\,|\gamma_{\boldsymbol{k}}|$. Again, electron spin is a spectator at this point, so each level is doubly degenerate in the absence of, *e.g.*, a Zeeman term[20].

It is convenient to write the crystal wavevector $\boldsymbol{k}$ as

$$\boldsymbol{k} \equiv \frac{\theta_1}{2\pi}\,\boldsymbol{b}_1 + \frac{\theta_2}{2\pi}\,\boldsymbol{b}_2 \tag{1.100}$$

in which case $\exp(i\boldsymbol{k}\cdot\boldsymbol{a}_{1,2}) = \exp(i\theta_{1,2})$.

1.4.3 Long wavelength Hamiltonian

When $u = v = w = 1$ and all the hopping amplitudes are equal to t, there are Dirac points at the two inequivalent zone corners $\boldsymbol{K}$ and $\boldsymbol{K}' = -\boldsymbol{K}$, where $\boldsymbol{K} = \frac{1}{3}(\boldsymbol{b}_1 + \boldsymbol{b}_2)$. At $\boldsymbol{k} = \boldsymbol{K}$, one has $\gamma_{\boldsymbol{K}} = 1 + e^{2\pi i/3} + e^{-2\pi i/3} = 0$, and similarly for $\boldsymbol{k} = \boldsymbol{K}'$. What happens when the hopping

[20] For a detailed description of the electronic properties of graphene, see A. H. Castro Neto *et al.*, *Rev. Mod. Phys.* **81**, 109 (2009).

amplitudes are not all the same? Having chosen the overall energy scale t, we may fix the sum $u + v + w = 3$ and choose to write

$$u = 1 + \tfrac{1}{2}\,\varepsilon_1 + \tfrac{1}{6}\,\varepsilon_2 \qquad , \qquad v = 1 - \tfrac{1}{2}\,\varepsilon_1 + \tfrac{1}{6}\,\varepsilon_2 \qquad , \qquad w = 1 - \tfrac{1}{3}\,\varepsilon_2 \quad . \tag{1.101}$$

We will write $\boldsymbol{k} = \boldsymbol{K} + \boldsymbol{q}$, with the deviation $\boldsymbol{q}$ from the Dirac point presumed to be small. Similarly, we presume that the dimensionless hopping anisotropies $\varepsilon_{1,2}$ are also small. Working to first order in smallness for each, we find

$$\begin{aligned}\gamma_{\boldsymbol{K}+\boldsymbol{q}} &= \big(1 + \tfrac{1}{2}\varepsilon_2 + \tfrac{1}{6}\,\varepsilon_3\big) + \big(1 - \tfrac{1}{2}\varepsilon_2 + \tfrac{1}{6}\,\varepsilon_3\big)\,e^{2\pi i/3}\,\big(1 + i\boldsymbol{q}\cdot\boldsymbol{a}_1 + \ldots\big) \\ &\qquad + \big(1 - \tfrac{1}{3}\,\varepsilon_3\big)\,e^{-2\pi i/3}\,\big(1 - i\boldsymbol{q}\cdot\boldsymbol{a}_2 + \ldots\big) \\ &= -\tfrac{\sqrt{3}}{2}\,(q_x - iq_y)\,a + \tfrac{1}{4}\,(3\varepsilon_1 + \varepsilon_2) - i\,\tfrac{\sqrt{3}}{4}\,(\varepsilon_1 - \varepsilon_2) + \ldots \\ &= -\tfrac{\sqrt{3}}{2}\,a\,\big(q_x - iq_y + Q_x - iQ_y\big) \quad ,\end{aligned} \tag{1.102}$$

where

$$Q_x \equiv -\frac{1}{2\sqrt{3}\,a}\,(3\varepsilon_1 + \varepsilon_2) \qquad , \qquad Q_y = \frac{1}{2a}\,(\varepsilon_2 - \varepsilon_1) \quad . \tag{1.103}$$

We may even allow $\varepsilon_{1,2}$ to vary slowly in space, in which case we must also impose the canonical commutation relations $\big[q_\alpha, x_\beta\big] = -i\,\delta_{\alpha\beta}$. Note that we may now read off $v_{\rm F} = \sqrt{3}\,ta/2\hbar$.

In the presence of an external gauge field, the prescription is the usual minimal coupling, *viz.*

$$\gamma_{\boldsymbol{K}+\boldsymbol{q}} = -\tfrac{\sqrt{3}}{2}\,a\left(q_x - iq_y + Q_x - iQ_y + \frac{e}{\hbar c}\,A_x - i\,\frac{e}{\hbar c}\,A_y\right) \quad , \tag{1.104}$$

Note that $q_x = -i\,\partial_x$ and $q_y = -i\,\partial_y$ are in fact differential operators for our purposes. Further defining

$$\theta = -\gamma_{\boldsymbol{K}+\boldsymbol{q}} = \tfrac{\sqrt{3}}{2}\,a\left[\left(q_x + Q_x + \frac{eA_x}{\hbar c}\right) - i\left(q_y + Q_y + \frac{eA_y}{\hbar c}\right)\right] \quad , \tag{1.105}$$

we derive

$$\big[\theta, \theta^\dagger\big] = \tfrac{3}{2}\,a^2\left(\partial_x Q_y - \partial_y Q_x + \frac{eB_z}{\hbar c}\right) \equiv \pm r^2 \quad . \tag{1.106}$$

We presume that r^2 is a constant, independent of space. If the sign on the RHS is positive, define $\theta \equiv r\alpha$, with $[\alpha, \alpha^\dagger] = 1$. Else, if negative, define $\theta \equiv r\alpha^\dagger$. In the former case, we have $\gamma_{\boldsymbol{K}+\boldsymbol{q}} = -r\alpha$ and in the latter $\gamma_{\boldsymbol{K}+\boldsymbol{q}} = -r\alpha^\dagger$. Thus,

$$H_{\boldsymbol{K}} = t\begin{pmatrix} 0 & \theta \\ \theta^\dagger & 0 \end{pmatrix} \qquad , \qquad H_{\boldsymbol{K},+} = rt\begin{pmatrix} 0 & \alpha \\ \alpha^\dagger & 0 \end{pmatrix} \qquad , \qquad H_{\boldsymbol{K},-} = rt\begin{pmatrix} 0 & \alpha^\dagger \\ \alpha & 0 \end{pmatrix} \quad . \tag{1.107}$$

Let's work out the eigenspectrum for H_+; the H_- case is equivalent since $H_- = \sigma^x\,H_+\,\sigma^x$. Let $\alpha\,|\,0\,\rangle = 0$ and $|\,n\,\rangle \equiv (\alpha^\dagger)^n\,|\,0\,\rangle/\sqrt{n!}$ with $n \in \mathbb{Z}_+$. Then it is easy to see that

$$H_+\,|\,\Psi_0\,\rangle = 0 \qquad , \qquad H_+\,|\,\Psi_{\pm,n}\,\rangle = \pm\sqrt{n}\,r\,t\,|\,\Psi_{\pm,n}\,\rangle \tag{1.108}$$

where

$$|\,\Psi_0\,\rangle = \begin{pmatrix} |\,0\,\rangle \\ 0 \end{pmatrix} \quad , \quad |\,\Psi_{\pm,n}\,\rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} |\,n\,\rangle \\ \pm|\,n-1\,\rangle \end{pmatrix} \quad . \tag{1.109}$$

The corresponding eigenstates of H_- are then given by $\sigma^x\,|\,\Psi_0\,\rangle$ and $\sigma^x\,|\,\Psi_{n,\pm}\,\rangle$. These are the Landau levels of the 2D Dirac Hamiltonian. As with the familiar case with continuum Landau levels from a ballistic dispersion, the Dirac Hamiltonian Landau levels are also massively degenerate. Note that the Landau level energy varies as $\sqrt{n}$ and not linearly in n as in the ballistic case!

1.4.4 The $\boldsymbol{K}'$ valley

Consider now the other inequivalent zone corner, located at $\boldsymbol{K}' = -\boldsymbol{K}$. We then have

$$\begin{aligned} \gamma_{-\boldsymbol{K}+\boldsymbol{q}} &= u + v\,e^{-2\pi i/3}\,e^{i\boldsymbol{q}\cdot\boldsymbol{a}_1} + w\,e^{2\pi i/3}\,e^{-i\boldsymbol{q}\cdot\boldsymbol{a}_2} \\ &= \tfrac{\sqrt{3}}{2}\,(q_x + iq_y)\,a + \tfrac{1}{4}(3\varepsilon_1 + \varepsilon_2) + i\,\tfrac{\sqrt{3}}{4}\,(\varepsilon_1 - \varepsilon_2) \\ &= \tfrac{\sqrt{3}}{2}\,a\,\big(q_x + iq_y - Q_x - iQ_y\big) \quad . \end{aligned} \tag{1.110}$$

Again we include the electromagnetic gauge field via minimal coupling,

$$\gamma_{-\boldsymbol{K}+\boldsymbol{q}} = \tfrac{\sqrt{3}}{2}\,a\left(q_x + iq_y - Q_x - iQ_y + \frac{e}{\hbar c}\,A_x + i\,\frac{e}{\hbar c}\,A_y\right) \quad , \tag{1.111}$$

and we define

$$\theta = -\gamma^*_{-\boldsymbol{K}+\boldsymbol{q}} = -\tfrac{\sqrt{3}}{2}\,a\left[\left(q_x - Q_x + \frac{eA_x}{\hbar c}\right) - i\left(q_y - Q_y + \frac{eA_y}{\hbar c}\right)\right] \quad , \tag{1.112}$$

we derive

$$\big[\theta, \theta^\dagger\big] = \tfrac{3}{2}\,a^2\left(-\,\partial_x Q_y + \partial_y Q_x + \frac{eB_z}{\hbar c}\right) \equiv \pm s^2 \quad . \tag{1.113}$$

The Hamiltonian in the $\boldsymbol{K}'$ valley is $H_{\boldsymbol{K}'} = t\begin{pmatrix} 0 & \theta^\dagger \\ \theta & 0 \end{pmatrix}$. Note that the magnetic flux has the same sign in Eqns. 1.106 and 1.113, but that the contribution from the hopping anisotropy is reversed. This is because magnetic flux breaks time reversal symmetry and hopping anisotropy does not, even though the spatially varying hopping anisotropy, which is due to strain, can generate Landau levels. What happens is that the Landau levels in the $\boldsymbol{K}$ valley are the time reverse states of the Landau levels in the $\boldsymbol{K}'$ valley.

1.4.5 Strain and pseudomagnetic fields

We now need a model for the distortions $\varepsilon_{1,2}$. Consider the case of *triaxial strain* where the local displacement at position (x, y) is given by $\boldsymbol{u}(x, y)$, where

$$u_x = 2\eta xy \quad , \quad u_y = \eta(x^2 - y^2) \quad . \tag{1.114}$$

Here η has dimensions of inverse length. Since hopping amplitudes typically vary exponentially with distance, we write, phenomenologically,

$$t_{\boldsymbol{\delta}} = t\,\exp\!\left(\frac{\delta_0 - |\boldsymbol{\delta}|}{\lambda}\right) \quad , \tag{1.115}$$

where $\delta_0 = a/\sqrt{3}$ is the unstrained nearest neighbor C-C bond length and λ is a constant with dimensions of length corresponding to the transverse extent of the atomic π orbital. Under strain, we have

$$\boldsymbol{\delta} \to \boldsymbol{\delta}' = \boldsymbol{\delta} + \boldsymbol{u}(\boldsymbol{R} + \boldsymbol{\delta}) - \boldsymbol{u}(\boldsymbol{R}) \quad , \tag{1.116}$$

and therefore to lowest nontrivial order

$$\begin{aligned} \delta'_x &= \delta_x + 2\eta\,(X\delta_y + Y\delta_x) \\ \delta'_y &= \delta_y + 2\eta\,(X\delta_x - Y\delta_y) \quad , \end{aligned} \tag{1.117}$$

with $\boldsymbol{R} = (X, Y)$. From these relations, we obtain the extensions

$$|\boldsymbol{\delta}'| - \delta_0 = \Delta_x(\boldsymbol{\delta})\,X + \Delta_y(\boldsymbol{\delta})\,Y \tag{1.118}$$

with

$$\Delta_x(\boldsymbol{\delta}) = 2\eta\delta_0 \cdot 2\hat{\delta}_x\hat{\delta}_y \qquad , \qquad \Delta_y(\boldsymbol{\delta}) = 2\eta\delta_0 \cdot \big(\hat{\delta}_x^2 - \hat{\delta}_y^2\big) \quad . \tag{1.119}$$

Thus,

$$\boldsymbol{\Delta}(\boldsymbol{\delta}_1) = -2\eta\delta_0\,\hat{\boldsymbol{y}} \quad , \quad \boldsymbol{\Delta}(\boldsymbol{\delta}_2) = 2\eta\delta_0\big(-\tfrac{\sqrt{3}}{2}\,\hat{\boldsymbol{x}} + \tfrac{1}{2}\,\hat{\boldsymbol{y}}\big) \quad , \quad \boldsymbol{\Delta}(\boldsymbol{\delta}_3) = 2\eta\delta_0\big(\tfrac{\sqrt{3}}{2}\,\hat{\boldsymbol{x}} + \tfrac{1}{2}\,\hat{\boldsymbol{y}}\big) \quad . \tag{1.120}$$

We now have

$$t(\boldsymbol{\delta}) = t\,\exp\!\left(\frac{\delta_0 - |\boldsymbol{\delta}|}{\lambda}\right) = t\big(1 - \boldsymbol{\Delta}(\boldsymbol{\delta}) \cdot \boldsymbol{R}/\lambda + \ldots\big) \quad , \tag{1.121}$$

which says

$$\begin{aligned} t_1(\boldsymbol{R}) &= \big(1 + \tfrac{1}{2}\,\varepsilon_1 + \tfrac{1}{6}\,\varepsilon_2\big)\,t = \big(1 + 2\eta\,\delta_0\,Y\lambda^{-1} + \ldots\big)\,t \\ t_2(\boldsymbol{R}) &= \big(1 - \tfrac{1}{2}\,\varepsilon_1 + \tfrac{1}{6}\,\varepsilon_2\big)\,t = \big(1 + \sqrt{3}\,\eta\,\delta_0\,X\lambda^{-1} - \eta\,\delta_0\,Y\lambda^{-1} + \ldots\big)\,t \\ t_3(\boldsymbol{R}) &= \big(1 - \tfrac{1}{3}\,\varepsilon_2\big)\,t = \big(1 - \sqrt{3}\,\eta\,\delta_0\,X\lambda^{-1} - \eta\,\delta_0\,Y\lambda^{-1} + \ldots\big)\,t \quad . \end{aligned} \tag{1.122}$$

Solving for $\varepsilon_{1,2}$, we obtain

$$\begin{aligned} \varepsilon_1 &= \sqrt{3}\,\eta\delta_0\,\big(-X + \sqrt{3}\,Y\big)\lambda^{-1} \\ \varepsilon_2 &= 3\eta\delta_0\,\big(\sqrt{3}\,X + Y\big)\lambda^{-1} \end{aligned} \tag{1.123}$$

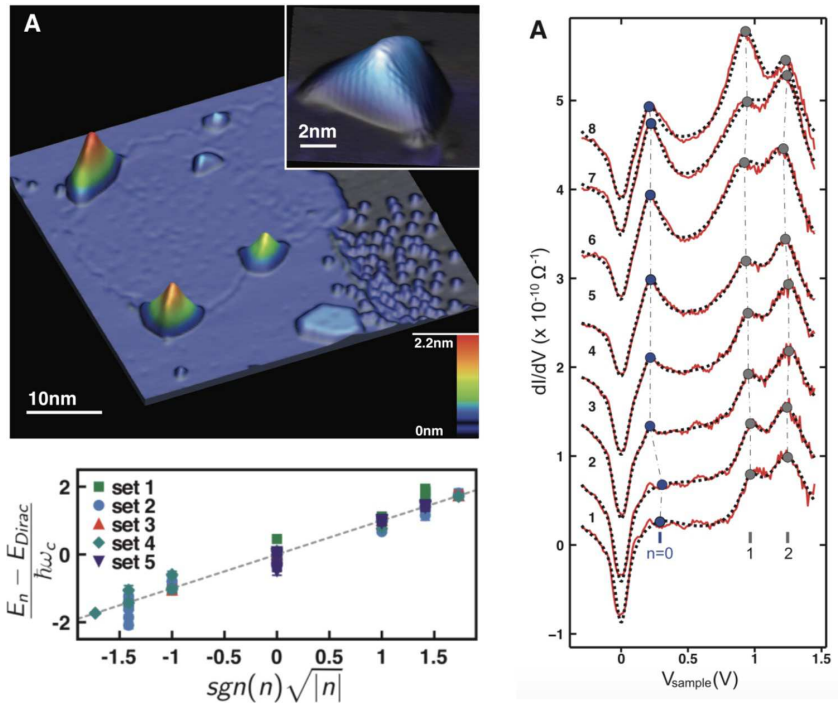


Figure 1.15: Strain-induced pseudo-Landau levels in graphene, reported by N. Levy *et al.*, *Science* **329**, 544 (2010). Upper left: STM images taken at $7.5\,\mathrm{K}$ showing a monolayer graphene patch on a Pt(111) surface. Several nanobubbles are visible. Right: STS spectra at $7.5\,\mathrm{K}$ showing locations of $n = 0$, $n = 1$, and $n = 2$ pseudo-Landau levels. Lower left: The energies of the pseudo-LLs are proportional to $|n|^{1/2}$, in agreement with the Dirac theory.

and therefore the effective strain gauge field components are

$$\begin{aligned} Q_x &= -\frac{1}{2\sqrt{3}\,a}\,(3\varepsilon_1 + \varepsilon_2) = -2\eta\,Y/\lambda \\ Q_y &= \frac{1}{2a}\,(\varepsilon_2 - \varepsilon_1) = 2\eta X/\lambda \quad . \end{aligned} \tag{1.124}$$

This is wonderful! The effective strain gauge field corresponds to a uniform fictitious magnetic field, with

$$\partial_x Q_y - \partial_y Q_x = 4\eta/\lambda \quad . \tag{1.125}$$

1.4.6 One-dimensional analog

There is a one-dimensional version of the strain gauge field. Consider a bipartite one-dimensional chain with alternating hoppings $t_1 \equiv ut$ and $t_2 \equiv vt$. The Hamiltonian at wavevector k is

$$H_k = -t \begin{pmatrix} 0 & \gamma_k \\ \gamma_k^* & 0 \end{pmatrix} \qquad \text{where} \qquad \gamma_k = u + v\,e^{ik} \quad . \tag{1.126}$$

The energy eigenvalues are $\pm t|\gamma_k|$. For uniform hopping $u = v = 1$, the gap collapses at $k = \pi$. Writing $u = 1 + \frac{1}{2}\varepsilon$ and $v = 1 - \frac{1}{2}\varepsilon$, and $k = \pi + q$, we obtain, to lowest order and in the long wavelength limit,

$$\pi^\dagger \equiv \gamma_{\pi+q} = -iq + \varepsilon = -\partial_x + \varepsilon(x) \quad . \tag{1.127}$$

Note $[\pi, \pi^\dagger] = 2\,\varepsilon'(x) \equiv \pm r^2$. Assuming r is both constant and real, we then define $\pi \equiv r\,\alpha$, and we once again have $H_+ = rt \begin{pmatrix} 0 & \alpha \\ \alpha^\dagger & 0 \end{pmatrix}$ and $H_- = rt \begin{pmatrix} 0 & \alpha^\dagger \\ \alpha & 0 \end{pmatrix}$, with the Dirac spectrum $E_0 = 0$ and $E_{\pm,n} = \sqrt{n}\,r\,t$.

1.5 An Electron on a Torus

1.5.1 Constraints of finite geometry

On an infinite plane, each Landau level is infinitely degenerate, with the number of states per Landau level per unit area given by $B/\phi_0 = 1/2\pi\ell^2$. As diagonalizing infinite-dimensional matrices is time-consuming, it is useful to impose a finite geometry, rendering each Landau level finite in size. Computational approaches have been exceedingly important in the development of the theory of the FQHE, often providing essential insights. In principle one could choose any orientable two-dimensional manifold as a base space, but the two simplest and most useful geometries have been the sphere and the torus. Use of the spherical geometry was pioneered by Duncan Haldane[21]. Here we will focus on the torus geometry[22]. One reason the torus is particularly useful is that it is not simply connected. Topologically, it is a product of two circles: $\mathbb{T}^2 = \mathbb{S}^1 \times \mathbb{S}^1$. This furnishes us with an opportunity to impose generalized periodic boundary conditions, which physically corresponds to threading each of the two circles with flux. Differentiating with respect to these fluxes can tell us about the Hall conductivity! So when it comes to finite geometries, to paraphrase Orwell: *sphere good, torus better!*

On the torus, we require that the area $\Omega = \hat{\boldsymbol{z}} \cdot \boldsymbol{L}_1 \times \boldsymbol{L}_2 = 2\pi\ell^2 N_\phi$ be quantized in units of $2\pi\ell^2$, where $N_\phi = B\Omega/\phi_0$ is the number of flux quanta, if we are to have $\big[\, t(\boldsymbol{L}_1)\,,\, t(\boldsymbol{L}_2)\,\big] = 0$. Then since $V(\boldsymbol{r} + \boldsymbol{L}_a) = V(\boldsymbol{r})$, we have that $\big\{\, H\,,\, t(\boldsymbol{L}_1)\,,\, t(\boldsymbol{L}_2)\,\big\}$ is a complete set of commuting observables. Since the $t(\boldsymbol{L}_{1,2})$ are unitary, the eigenstates $|\,\psi_\alpha\,\rangle$ of H may be chosen to satisfy

$$t(\boldsymbol{L}_a)\,|\,\psi_\alpha\,\rangle = e^{i\theta_a}\,|\,\psi_\alpha\,\rangle \tag{1.128}$$

for all $\alpha \in \big\{1, \ldots, N_\phi\big\}$. The dimension of the Hilbert space is N_ϕ, as we shall see.

We will consider potentials $V(\boldsymbol{r})$ which are periodic on the torus. This is equivalent to a periodic potential, and any reciprocal lattice vector $\boldsymbol{G}$ as

$$\boldsymbol{G} = n_1\,\boldsymbol{b}_1 + n_2\,\boldsymbol{b}_2 \quad , \tag{1.129}$$

[21] See F. D. M. Haldane, *Phys. Rev. Lett.* **31**, 605 (1983).
[22] Also largely pioneered by Haldane.

where

$$\boldsymbol{b}_1 = \frac{2\pi}{\Omega}\,\boldsymbol{L}_2 \times \hat{\boldsymbol{z}} \qquad , \qquad \boldsymbol{b}_2 = \frac{2\pi}{\Omega}\,\hat{\boldsymbol{z}} \times \boldsymbol{L}_1 \quad , \tag{1.130}$$

with integer $n_{1,2}$, where $\Omega = \hat{\boldsymbol{z}} \cdot \boldsymbol{L}_1 \times \boldsymbol{L}_2 = 2\pi\ell^2 N_\phi$ is the area of the torus. With the above definitions of the primitive reciprocal lattice vectors $\boldsymbol{b}_{1,2}$, we have $\boldsymbol{b}_a \cdot \boldsymbol{L}_{a'} = 2\pi\delta_{aa'}$. Thus,

$$\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G} = \frac{2\pi\ell^2}{\Omega}\left(n_1\,\boldsymbol{L}_2 - n_2\,\boldsymbol{L}_1\right) \quad . \tag{1.131}$$

Thus,

$$\left(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G}\right) \times \boldsymbol{L}_a = -2\pi\ell^2\, n_a\, \hat{\boldsymbol{z}} \quad , \tag{1.132}$$

and therefore $\left[\, t\big(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G}\big)\,,\, t(\boldsymbol{L}_a)\,\right] = 0$ for $a = 1, 2$.

For any vector $\boldsymbol{Q}$, we define its complexification as $Q \equiv Q_x + iQ_y$. The complexified elementary reciprocal lattice vectors are then

$$b_1 = b_{1,x} + ib_{1,y} = -\frac{2\pi i}{\Omega} L_2 \qquad , \qquad b_2 = b_{2,x} + ib_{2,y} = \frac{2\pi i}{\Omega} L_1 \quad . \tag{1.133}$$

The modular parameter $\tau = \tau_1 + i\tau_2$ is defined as the complex ratio

$$\tau \equiv \frac{L_2}{L_1} = \frac{L_{2,x} + iL_{2,y}}{L_{1,x} + iL_{1,y}} \quad . \tag{1.134}$$

For a general reciprocal lattice vector $\boldsymbol{G} = n_1\boldsymbol{b}_1 + n_2\boldsymbol{b}_2$, then, we have

$$G = G_x + iG_y = \frac{2\pi i}{\Omega}\, L_1\left(n_2 - n_1\tau\right) \quad . \tag{1.135}$$

The unit cell area is then

$$\Omega = 2\pi\ell^2 N = \mathsf{Im}\left(\bar{L}_1 L_2\right) = |L_1|^2\,\tau_2 \quad . \tag{1.136}$$

Then we have

$$\tfrac{1}{4}\boldsymbol{G}^2\ell^2 = \tfrac{1}{4}|G|^2\ell^2 = \frac{\pi}{2N_\phi\tau_2}\Big((n_1\tau_1 - n_2)^2 + n_1^2\,\tau_2^2\Big) \quad . \tag{1.137}$$

1.5.2 Lowest Landau level Hamiltonian

The potential may be written in terms of its Fourier components, *viz.*

$$\begin{aligned} V(\boldsymbol{r}) &= \sum_{\boldsymbol{G}} V_{\boldsymbol{G}}\, e^{i\boldsymbol{G}\cdot\boldsymbol{r}} = \sum_{\boldsymbol{G}} V_{\boldsymbol{G}}\, e^{i(G\bar{z} + \bar{G}z)/2} \\ &= \sum_{\boldsymbol{G}} V_{\boldsymbol{G}}\, e^{\ell^2(G\pi^\dagger - \bar{G}\pi)/2\hbar}\, e^{\ell^2(\bar{G}\kappa - G\kappa^\dagger)/2\hbar} \\ &= \sum_{\boldsymbol{G}} V_{\boldsymbol{G}}\, e^{-\boldsymbol{G}^2\ell^2/4}\, e^{\ell^2 G\pi^\dagger/2\hbar} e^{-\ell^2\bar{G}\pi/2\hbar}\, t\big(\ell^2\hat{\boldsymbol{z}} \times \boldsymbol{G}\big) \quad . \end{aligned} \tag{1.138}$$

If we project onto the lowest Landau level, we obtain

$$\widetilde{V} = \mathrm{P}_0\, V(\boldsymbol{r})\, \mathrm{P}_0 = \sum_{\boldsymbol{G}} V_{\boldsymbol{G}}\, e^{-\boldsymbol{G}^2\ell^2/4}\, t\big(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G}\big) \quad . \tag{1.139}$$

Define the unitary operators

$$t_1 \equiv t\big(\boldsymbol{L}_1/N_\phi\big) \qquad , \qquad t_2 \equiv t\big(\boldsymbol{L}_2/N_\phi\big) \quad . \tag{1.140}$$

Then it is easy to show

$$t_1\, t_2 = e^{2\pi i/N_\phi}\, t_2\, t_1 \quad . \tag{1.141}$$

Furthermore, we have

$$t\big(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G}\big) = t\left(\frac{n_1\, \boldsymbol{L}_2}{N_\phi} - \frac{n_2\, \boldsymbol{L}_1}{N_\phi}\right) = e^{-i\pi n_1 n_2/N_\phi}\, t_2^{n_1}\, t_1^{-n_2} \quad . \tag{1.142}$$

We can define an N_ϕ-element basis $\big\{ |\, k\,\rangle \big\}$ which satisfies the following:

$$\begin{aligned} t_1\, |\, k\,\rangle &= e^{i\theta_1/N_\phi}\, |\, k-1\,\rangle \\ t_2\, |\, k\,\rangle &= e^{i\theta_2/N_\phi}\, e^{2\pi i k/N_\phi}\, |\, k\,\rangle \quad , \end{aligned} \tag{1.143}$$

with $|\, k + N_\phi\,\rangle \equiv |\, k\,\rangle$. Note that $t_1\, t_2\, |\, k\,\rangle = e^{2\pi i/N_\phi}\, t_2\, t_1\, |\, k\,\rangle$ for all k, and furthermore that $t(\boldsymbol{L}_a)\, |\, k\,\rangle = t_a^{N_\phi}\, |\, k\,\rangle = e^{i\theta_a}\, |\, k\,\rangle$ for all allowed a and k. Thus,

$$\begin{aligned} \langle\, k\,|\, t\big(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{G}\big)\,|\, k'\,\rangle &= e^{-i\pi n_1 n_2/N_\phi}\, e^{in_1\theta_2/N_\phi}\, e^{2\pi i k n_1/N_\phi}\, e^{-in_2\theta_1/N_\phi}\, \langle\, k\,|\, k' + n_2\,\rangle \\ &= e^{-i\pi n_1 n_2/N_\phi}\, e^{i(n_1\theta_2 - n_2\theta_1)/N_\phi}\, e^{2\pi i k n_1/N_\phi}\, \tilde{\delta}_{k,k'+n_2} \quad , \end{aligned} \tag{1.144}$$

where $\tilde{\delta}_{k,l} \equiv \delta_{k,l \bmod N_\phi}$. Thus, the Hamiltonian for our system is

$$H_{kk'}(\theta_1,\theta_2) = \sum_{n_1,n_2} V_{n_1,n_2}\, e^{-\pi\left[(n_1\tau_1 - n_2)^2 + n_1^2\tau_2^2\right]/2N_\phi\tau_2}\, e^{-i\pi n_1 n_2/N_\phi}\, e^{i(n_1\theta_2 - n_2\theta_1)/N_\phi}\, e^{2\pi i n_1 k/N_\phi}\, \tilde{\delta}_{k,k'+n_2} \quad . \tag{1.145}$$

Checking that the Hamiltonian is Hermitian, we have

$$\begin{aligned} H^*_{kk'} &= \sum_{n_1,n_2} V^*_{n_1,n_2}\, e^{-\pi(n_1\tau_1 - n_2)^2/2N_\phi\tau_2}\, e^{-\pi n_1^2\tau_2/2N_\phi}\, e^{i\pi n_1 n_2/N_\phi}\, e^{-i(n_1\theta_2 - n_2\theta_1)/N_\phi}\, e^{-2\pi i n_1 k/N_\phi}\, \tilde{\delta}_{k,k'+n_2} \\ &= \sum_{n_1,n_2} V_{n_1,n_2}\, e^{-\pi(n_1\tau_1 - n_2)^2/2N_\phi\tau_2}\, e^{-\pi n_1^2\tau_2/2N_\phi}\, e^{i\pi n_1 n_2/N_\phi}\, e^{i(n_1\theta_2 - n_2\theta_1)/N_\phi}\, e^{2\pi i n_1 (k'-n_2)/N_\phi}\, \tilde{\delta}_{k',k+n_2} \\ &= \sum_{n_1,n_2} V_{n_1,n_2}\, e^{-\pi(n_1\tau_1 - n_2)^2/2N_\phi\tau_2}\, e^{-\pi n_1^2\tau_2/2N_\phi}\, e^{-i\pi n_1 n_2/N_\phi}\, e^{i(n_1\theta_2 - n_2\theta_1)/N_\phi}\, e^{2\pi i n_1 k'/N_\phi}\, \tilde{\delta}_{k',k+n_2} = H_{k'k} \quad , \end{aligned}$$

where we have used $V^*_{n_1,n_2} = V_{-n_1,-n_2}$ and then replaced $n_{1,2}$ with $-n_{1,2}$ in the second line.

As an example, consider a case where $V(\boldsymbol{r}) = V_{1,1}\, e^{i(\boldsymbol{b}_1+\boldsymbol{b}_2)\cdot\boldsymbol{r}} + V^*_{1,1}\, e^{-i(\boldsymbol{b}_1+\boldsymbol{b}_2)\cdot\boldsymbol{r}}$ and $N_\phi = 3$. We then find

$$H = \begin{pmatrix} 0 & -V^*_{1,1}\, e^{i\alpha} & V_{1,1}\, e^{i\pi/3}\, e^{-i\alpha} \\ -V_{1,1}\, e^{-i\alpha} & 0 & V^*_{1,1}\, e^{i\pi/3}\, e^{i\alpha} \\ V^*_{1,1}\, e^{-i\pi/3}\, e^{i\alpha} & V_{1,1}\, e^{-i\pi/3}\, e^{-i\alpha} & 0 \end{pmatrix} e^{-\pi\left[(\tau_1-1)^2+\tau_2^2\right]/6\tau_2} \quad , \tag{1.146}$$

where $\alpha \equiv \frac{1}{3}(\theta_1 - \theta_2)$. For example, to compute $H_{kk'}$ with $k = 1$ and $k' = 2$, we need $n_2 = -1$ to satisfy the Kronecker delta in eqn. 1.145, and therefore $n_1 = -1$ as well, corresponding to $V_{-1,-1} = V^*_{1,1}$. Working out the phase of the matrix element, we then have $e^{-i\pi/3}\, e^{i(\theta_1-\theta_2)/3}\, e^{-2\pi i/3} = -e^{i\alpha}$. For $k = 1$ and $k' = 3$, we need $n_2 = 1$ in order to satisfy $k = k' + 1 \bmod N$. Thus $n_1 = 1$ as well, corresponding to $V_{1,1}$, and the phase of the matrix element is $e^{-i\pi/3}\, e^{i(\theta_2-\theta_1)/3}\, e^{2\pi i/3} = e^{i\pi/3}\, e^{-i\alpha}$.

A detailed discussion of the LLL wavefunctions on the torus is given in §1.8 below.

1.6 Lattice Models and Hofstadter's Butterfly

1.6.1 Tight binding with $B = 0$

As you may have heard, solids are composed of atoms[23]. Suppose we have an orthonormal set of orbitals $|\, a\boldsymbol{R}\,\rangle$, where a labels the orbital and $\boldsymbol{R}$ denotes a Bravais lattice site. The label a may refer to different orbitals associated with the atom at $\boldsymbol{R}$, or it may label orbitals on other atoms in the unit cell defined by $\boldsymbol{R}$. The most general tight binding Hamiltonian we can write is

$$H = \sum_{\boldsymbol{R},\boldsymbol{R}'} \sum_{a,a'} H_{aa'}(\boldsymbol{R}-\boldsymbol{R}')\, |\, a\boldsymbol{R}\,\rangle\langle\, a'\boldsymbol{R}'\,| \quad , \tag{1.147}$$

where $H_{aa'}(\boldsymbol{R}-\boldsymbol{R}') = H^*_{a'a}(\boldsymbol{R}'-\boldsymbol{R}) = \langle\, a, \boldsymbol{R}\,|\, H\,|\, a', \boldsymbol{R}'\,\rangle$ is the Hamiltonian matrix, whose rows and columns are indexed by a composite index combining both the unit cell label $\boldsymbol{R}$ and the orbital label a. When $\boldsymbol{R} = \boldsymbol{R}'$ and $a = a'$, the term $H_{aa}(0) = \varepsilon_a$ is the energy of a single electron in an isolated a orbital. For all other cases, $H_{aa'}(\boldsymbol{R}-\boldsymbol{R}') = -t_{aa'}(\boldsymbol{R}-\boldsymbol{R}')$ is the hopping integral between the a orbital in unit cell $\boldsymbol{R}$ and the a' orbital in unit cell $\boldsymbol{R}'$. Let's write an eigenstate $|\,\psi\,\rangle$ as

$$|\,\psi\,\rangle = \sum_{\boldsymbol{R}} \sum_{a} \psi_{a\boldsymbol{R}}\, |\, a\boldsymbol{R}\,\rangle \quad . \tag{1.148}$$

Applying the Hamiltonian to $|\,\psi\,\rangle$, we obtain the coupled equations

$$\sum_{\boldsymbol{R},\boldsymbol{R}'} \sum_{a,a'} H_{aa'}(\boldsymbol{R}-\boldsymbol{R}')\, \psi_{a'\boldsymbol{R}'}\, |\, a\boldsymbol{R}\,\rangle = E \sum_{\boldsymbol{R}} \sum_{a} \psi_{a\boldsymbol{R}}\, |\, a\boldsymbol{R}\,\rangle \quad . \tag{1.149}$$

[23] A somewhat more nuanced description: solids are composed of ions and electrons.

Since the $|\, a\boldsymbol{R}\,\rangle$ basis is complete, we must have that the coefficients of $|\, a\boldsymbol{R}\,\rangle$ on each side agree. Therefore,

$$\sum_{\boldsymbol{R}'}\sum_{a'} H_{aa'}(\boldsymbol{R}-\boldsymbol{R}')\,\psi_{a'\boldsymbol{R}'} = E\,\psi_{a\boldsymbol{R}} \quad . \tag{1.150}$$

Bloch's theorem

We now use Bloch's theorem, which says that each eigenstate may be labeled by a wavevector $\boldsymbol{k}$, with $\psi_{a\boldsymbol{R}} = \frac{1}{\sqrt{N}}\,u_a(\boldsymbol{k})\,e^{i\boldsymbol{k}\cdot\boldsymbol{R}}$. The $N^{-1/2}$ prefactor is a normalization term. Multiplying each side by $e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}$, we have

$$\sum_{a'}\left(\sum_{\boldsymbol{R}'} H_{aa'}(\boldsymbol{R}-\boldsymbol{R}')\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{R}-\boldsymbol{R}')}\right) u_{a'\boldsymbol{k}} = E(\boldsymbol{k})\,u_{a\boldsymbol{k}} \quad , \tag{1.151}$$

which may be written as

$$\sum_{a'} \hat{H}_{aa'}(\boldsymbol{k})\,u_{a'}(\boldsymbol{k}) = E(\boldsymbol{k})\,u_a(\boldsymbol{k}) \quad , \tag{1.152}$$

where

$$\hat{H}_{aa'}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} H_{aa'}(\boldsymbol{R})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{R}} \quad . \tag{1.153}$$

Thus, for each crystal wavevector $\boldsymbol{k}$, the $u_{a\boldsymbol{k}}$ are the eigenfunctions of the $r\times r$ Hermitian matrix $\hat{H}_{aa'}(\boldsymbol{k})$. The energy eigenvalues at wavevector $\boldsymbol{k}$ are given by $\mathrm{spec}\big(\hat{H}(\boldsymbol{k})\big)$, *i.e.* by the set of eigenvalues of the matrix $\hat{H}(\boldsymbol{k})$. There are r such solutions (some of which may be degenerate), which we distinguish with a band index n, and we denote $u_{na}(\boldsymbol{k})$ and $E_n(\boldsymbol{k})$ as the corresponding eigenvectors and eigenvalues. We sometimes will use the definition $\hat{t}_{aa'}(\boldsymbol{k}) \equiv -\hat{H}_{aa'}(\boldsymbol{k})$ for the matrix of hopping integrals.

In Eqn. 1.147, $\boldsymbol{R}$ and $\boldsymbol{R}'$ label Bravais lattice sites, while a and a' label orbitals. We stress that these orbitals don't necessarily have to be located on the same ion. We should think of $\boldsymbol{R}$ and $\boldsymbol{R}'$ labeling *unit cells*, each of which is indeed associated with a Bravais lattice site. For example, in the case of graphene, $|\, a\boldsymbol{R}\,\rangle$ represents an orbital on the a sublattice in unit cell $\boldsymbol{R}$. The eigenvalue equation may be written

$$\hat{H}_{aa'}(\boldsymbol{k})\,u_{na'}(\boldsymbol{k}) = E_n(\boldsymbol{k})\,u_{na}(\boldsymbol{k}) \quad , \tag{1.154}$$

where n is the band index. The function $u_{na}(\boldsymbol{k})$ is the *internal wavefunction* within a given cell, and corresponds to the cell function $u_{n\boldsymbol{k}}(\boldsymbol{r})$ in the continuum, with $a \leftrightarrow (\boldsymbol{r}-\boldsymbol{R})$ labeling a position within each unit cell. The full Bloch state may then be written

$$|\,\psi_{n\boldsymbol{k}}\,\rangle = |\,\boldsymbol{k}\,\rangle \otimes |\,u_{n\boldsymbol{k}}\,\rangle \quad , \tag{1.155}$$

so that

$$\begin{aligned}\psi_{n\boldsymbol{k}}(\boldsymbol{R},a) &= \Big(\langle\,\boldsymbol{R}\,|\otimes\langle\,a\,|\Big)\Big(|\,\boldsymbol{k}\,\rangle\otimes|\,u_{n\boldsymbol{k}}\,\rangle\Big)\\ &= \langle\,\boldsymbol{R}\,|\,\boldsymbol{k}\,\rangle\,\langle\,a\,|\,u_{n\boldsymbol{k}}\,\rangle = \frac{1}{\sqrt{N}}\,e^{i\boldsymbol{k}\cdot\boldsymbol{R}}\,u_{na}(\boldsymbol{k}) \quad .\end{aligned} \tag{1.156}$$

Here we have chosen a normalization $\sum_a \left|u_{na}(\boldsymbol{k})\right|^2 = 1$ within each unit cell, which entails the overall normalization $\sum_{\boldsymbol{R},a}\left|\psi_{n\boldsymbol{k}}(\boldsymbol{R},a)\right|^2 = 1$.

1.6.2 Go flux yourself : how to add magnetic fields

To simplify matters, we consider only s-orbitals on two-dimensional lattices. The general tight-binding Hamiltonian is written

$$H = -\sum_{\boldsymbol{r}<\boldsymbol{r}'}\Big(t_{\boldsymbol{r}\boldsymbol{r}'}\,c^\dagger_{\boldsymbol{r}}\,c_{\boldsymbol{r}'} + t^*_{\boldsymbol{r}\boldsymbol{r}'}\,c^\dagger_{\boldsymbol{r}'}\,c_{\boldsymbol{r}}\Big) \quad , \tag{1.157}$$

where the notation $\boldsymbol{r} < \boldsymbol{r}'$ means that each pair $(\boldsymbol{r},\boldsymbol{r}')$ summed only once. We may write $t_{\boldsymbol{r}\boldsymbol{r}'} = t^*_{\boldsymbol{r}'\boldsymbol{r}} = |t_{\boldsymbol{r}\boldsymbol{r}'}|\exp(iA_{\boldsymbol{r}\boldsymbol{r}'})$, where $A_{\boldsymbol{r}\boldsymbol{r}'}$ is a gauge field living on the links of the lattice. Let p denote a plaquette on the lattice. Then the dimensionless flux ϕ_p (in units of $\hbar c/e$) through plaquette p is

$$\phi_p = \sum_{\langle\boldsymbol{r}\boldsymbol{r}'\rangle\in\partial p} A_{\boldsymbol{r}\boldsymbol{r}'} \quad , \tag{1.158}$$

where the sum is taken in a counterclockwise fashion along the links on the boundary of p. The tight-binding Hamiltonian exhibits a *gauge invariance* under

$$\begin{aligned}c_{\boldsymbol{r}} &\to e^{i\alpha_{\boldsymbol{r}}}\,c_{\boldsymbol{r}}\\ t_{\boldsymbol{r}\boldsymbol{r}'} &\to e^{i(\alpha_{\boldsymbol{r}}-\alpha_{\boldsymbol{r}'})}\,t_{\boldsymbol{r}\boldsymbol{r}'} \quad .\end{aligned} \tag{1.159}$$

Consider now the case of the square lattice. It is clear that any configuration of the $A_{\boldsymbol{r}\boldsymbol{r}'}$ which is periodic in the structural unit cell, *i.e.* under translations by elementary direct lattice vectors, must correspond to $\phi_p = 0$ for every plaquette p[24]. This is because the phase $A_{\boldsymbol{r}\boldsymbol{r}'}$ is associated with the *directed link* from $\boldsymbol{r}$ to $\boldsymbol{r}'$, and parallel links on opposite sides of the elementary square plaquette will yield equal and opposite values of $A_{\boldsymbol{r}\boldsymbol{r}'}$ because they are traversed in opposite directions. *In order to describe nonzero flux per plaquette, the configuration of the lattice gauge field*

[24]More precisely, if $A_{\boldsymbol{r}\boldsymbol{r}'}$ is periodic in the structural unit cell, then each structural unit cell is congruent to a zero flux state. However, it may be that a structural cell is comprised of more than one elementary plaquette, as is the case with the triangular lattice (each structural cell consists of two triangles), or that there are closed loops which don't correspond to a structural unit cell due to further neighbor hoppings. In such cases, there may be closed loops on the lattice whose flux is not congruent to zero. See §1.6.3 for some examples.

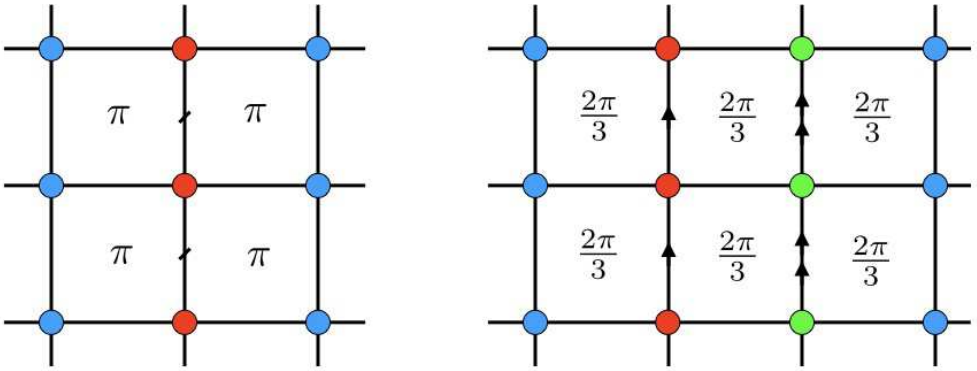


Figure 1.16: Gauges for the square lattice Hofstadter model. Left: $\phi = \pi$ case. $t_{\boldsymbol{r}\boldsymbol{r}'} = t$ on all links except those with slashes, where $t_{\boldsymbol{r}\boldsymbol{r}'} = -t$. Right: $\phi = \frac{2}{3}\,\pi$. Each arrow corresponds to a factor of $\exp(2\pi i/3)$.

$A_{\boldsymbol{r}\boldsymbol{r}'}$ must break lattice translational symmetry[25]. Consider the case where $\phi = \pi$ in each plaquette. A configuration for the gauge field $A_{\boldsymbol{r}\boldsymbol{r}'}$ yielding this flux distribution is shown in the left panel of Fig. 1.16. All links have $A_{\boldsymbol{r}\boldsymbol{r}'} = 0$, hence $t_{\boldsymbol{r}\boldsymbol{r}'} = t\exp(iA_{\boldsymbol{r}\boldsymbol{r}'}) = t$, except for the links depicted with slashes, for which $A_{\boldsymbol{r}\boldsymbol{r}'} = \pi$ and $t_{\boldsymbol{r}\boldsymbol{r}'} = -t$. The *magnetic unit cell* is now a 2×1 block consisting of one cell from each sublattice (blue and red). We call this a magnetic unit cell to distinguish it from the *structural unit cell* of the underlying square lattice. The *structural* Bravais lattice is square, with elementary direct lattice vectors are $\boldsymbol{a}_1 = a\hat{\boldsymbol{x}}$ and $\boldsymbol{a}_2 = a\hat{\boldsymbol{y}}$. But the *magnetic* Bravais lattice is rectangular, with elementary RLVs $\boldsymbol{a}_1 = 2a\hat{\boldsymbol{x}}$ and $\boldsymbol{a}_2 = a\hat{\boldsymbol{y}}$. From Bloch's theorem, the phase of the wavefunction varies by $\exp(i\boldsymbol{k}\cdot\boldsymbol{a}_1) \equiv \exp(i\theta_1)$ across the unit cell in the x-direction, and by $\exp(i\boldsymbol{k}\cdot\boldsymbol{a}_2) \equiv \exp(i\theta_2)$ in the y-direction. The Hamiltonian is

$$\hat{H}(\boldsymbol{\theta}) = -t\begin{pmatrix} 2\cos\theta_2 & 1+e^{-i\theta_1} \\ 1+e^{i\theta_1} & -2\cos\theta_2 \end{pmatrix} \tag{1.160}$$

The energy eigenvalues are $E_\pm(\boldsymbol{\theta}) = \pm 2t\sqrt{\cos^2(\frac{1}{2}\theta_1) + \cos^2\theta_2}$. The band gap collapses at two points: $(\theta_1, \theta_2) = (\pi, \pm\frac{1}{2}\pi)$. Writing $(\theta_1, \theta_2) = (\pi + \delta_1\,,\, \pm\frac{1}{2}\pi + \delta_2)$, we find

$$E_\pm(\boldsymbol{\theta}) = \pm 2t\sqrt{\sin^2(\tfrac{1}{2}\delta_1) + \sin^2\delta_2} = \pm 2ta\,\sqrt{\tfrac{1}{4}q_1^2 + q_2^2} \quad , \tag{1.161}$$

which is a Dirac cone! Thus, the dispersion for the square lattice π flux model has two Dirac points. Here, $\boldsymbol{q} = \boldsymbol{k} - \boldsymbol{k}_{\rm D}$ is the wavevector measured from either Dirac point.

The π flux state is time-reversal symmetric, since under time reversal we have $\exp(iA_{\boldsymbol{r}\boldsymbol{r}'}) \to \exp(-iA_{\boldsymbol{r}\boldsymbol{r}'})$, hence $\phi_p \to -\phi_p$. But flux is only defined modulo 2π, hence $\pi \to -\pi \cong \pi$ yields the same flux configuration.

A more interesting state of affairs pertains for the case $\phi = \frac{2}{3}\pi$, for which a valid gauge configuration $A_{\boldsymbol{r}\boldsymbol{r}'}$ is shown in the right panel of Fig. 1.16. Now there are three sites per unit cell: A

[25] By "nonzero" flux, we mean $\phi \bmod 2\pi \neq 0$.

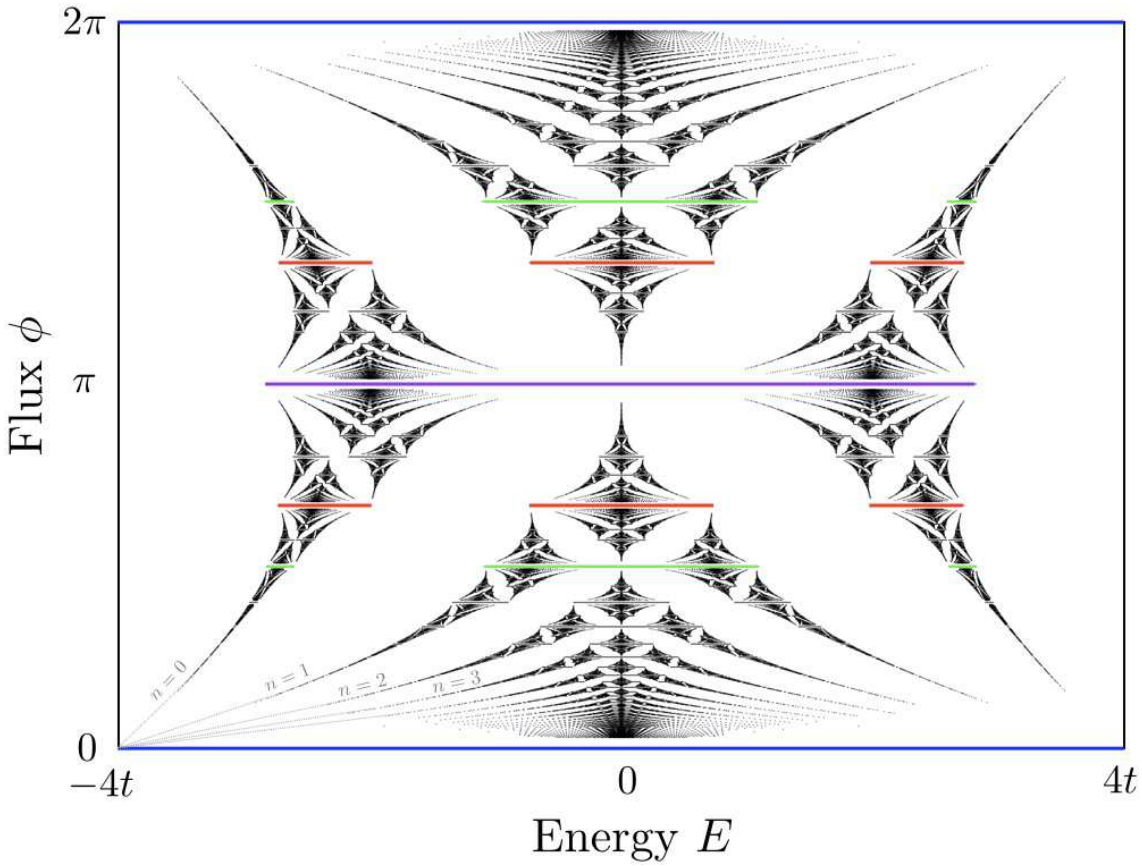


Figure 1.17: Magnetic subbands for the square lattice Hofstadter model for flux per plaquette $\phi \in [0, 2\pi]$. Blue bands at $\phi = 0$ and $\phi = 2\pi$ have the full bandwidth $W = 8t$. At $\phi = \pi$ (purple), there are two subbands with $E_- \in [-2\sqrt{2}t, 0]$ and $E_+ \in [0, 2\sqrt{2}t]$ which touch at $E = 0$. Similarly, at $\phi = \pm\frac{1}{2}\pi$ (green), there are four subbands, with the central two bands!bagain touching at $E = 0$. At $\phi = \pm\frac{2}{3}\pi$ (red), there are three subbands. Continuum Landau levels are shown radiating from the lower left corner.

(blue), B (red), and C (green). The Bloch phase accrued across the magnetic unit cell in the $\pm\hat{\boldsymbol{x}}$ direction is $e^{\pm i\theta_1}$, and in the $\pm\hat{\boldsymbol{y}}$ direction is $e^{\pm i\theta_2}$. Thus

$$\hat{H}(\boldsymbol{\theta}) = -t \begin{pmatrix} 2\cos\theta_2 & 1 & e^{-i\theta_1} \\ 1 & 2\cos(\theta_2 + \frac{2\pi}{3}) & 1 \\ e^{i\theta_1} & 1 & 2\cos(\theta_2 + \frac{4\pi}{3}) \end{pmatrix} \quad . \tag{1.162}$$

The general case where the flux per structural unit cell is $\phi = 2\pi p/q$ is known as the Hofstadter model[26] In this case, the magnetic unit cell is a $q \times 1$ block, and the resulting $q \times q$ Hamiltonian

[26]See D. R. Hofstadter, *Phys. Rev. B* **14**, 2239 (1976).

is given by

$$\hat{H}(\boldsymbol{\theta}) = -t \begin{pmatrix} 2\cos\theta_2 & 1 & 0 & \cdots & 0 & e^{-i\theta_1} \\ 1 & 2\cos\left(\theta_2 + \frac{2\pi p}{q}\right) & 1 & & & 0 \\ 0 & 1 & 2\cos\left(\theta_2 + \frac{4\pi p}{q}\right) & 1 & & \vdots \\ \vdots & 0 & 1 & \ddots & & \vdots \\ & & & & & \\ 0 & & & & & 1 \\ e^{i\theta_1} & 0 & & \cdots & 1 & 2\cos\left(\theta_2 + \frac{2\pi(q-1)p}{q}\right) \end{pmatrix} \quad . \tag{1.163}$$

There are thus q magnetic subbands. Note that

$$H(\theta_1, \theta_2 + \tfrac{2\pi p}{q}) = XUH(\theta_1, \theta_2)\, U^\dagger X^\dagger \quad , \tag{1.164}$$

where $X_{ij} = \delta_{i,j+1 \bmod q}$ and $U = \mathsf{diag}(1, e^{i\theta_1}, \cdots, e^{i\theta_1})$. Thus,

$$\mathsf{spec}\, H(\theta_1, \theta_2 + \tfrac{2\pi p}{q}) = \mathsf{spec}\, H(\theta_1, \theta_2) \quad , \tag{1.165}$$

as we saw explicitly in the $q = 2$ case above. A plot of the magnetic subbands in (E, ϕ) space, known as *Hofstadter's butterfly*, is shown in Fig. 1.17.

In the limit where the denominator q of the flux $\phi = 2\pi p/q$ is large (for fixed p), the flux per cell is very small. We then expect to recover the continuum Landau level spectrum $E_n = (n+\frac{1}{2})\hbar\omega_{\rm c}$. To express this in terms of the flux ϕ, note that the $B = 0$ dispersion is

$$E(\boldsymbol{k}) = -2t\cos(k_x a) - 2t\cos(k_y a) = -4t + t\boldsymbol{k}^2 a^2 + \ldots \quad , \tag{1.166}$$

which allows us to identify the effective mass m from the coefficient of the $\boldsymbol{k}^2$ term, with the result $m = \hbar^2/2ta^2$. The magnetic field is the flux per unit area, hence $B = \phi\hbar c/ea^2$. Thus,

$$\hbar\omega_{\rm c} = \frac{\hbar eB}{mc} = \frac{\hbar e}{c} \times \frac{\phi\,\hbar c}{ea^2} \times \frac{2ta^2}{\hbar^2} = 2\phi\, t \quad . \tag{1.167}$$

This describes the corners of the Hofstadter butterfly in Fig. 1.17, where continuum Landau levels radiate outward from the energies $\pm 4t$ according to

$$E_n(\phi) = \pm\Big(4t - (2n+1)\,\phi\, t\Big) \qquad \text{and} \qquad E_n(\phi) = \pm\Big(4t - (2n+1)\,(2\pi - \phi)\, t\Big) \quad , \tag{1.168}$$

for $\phi \ll 1$.

1.6.3 Unit cells with zero net flux

As mentioned in a footnote above, it is not quite true that a lattice gauge field $A_{rr'}$ which is periodic in the underlying Bravais lattice unit cell leads to zero net flux in every plaquette or

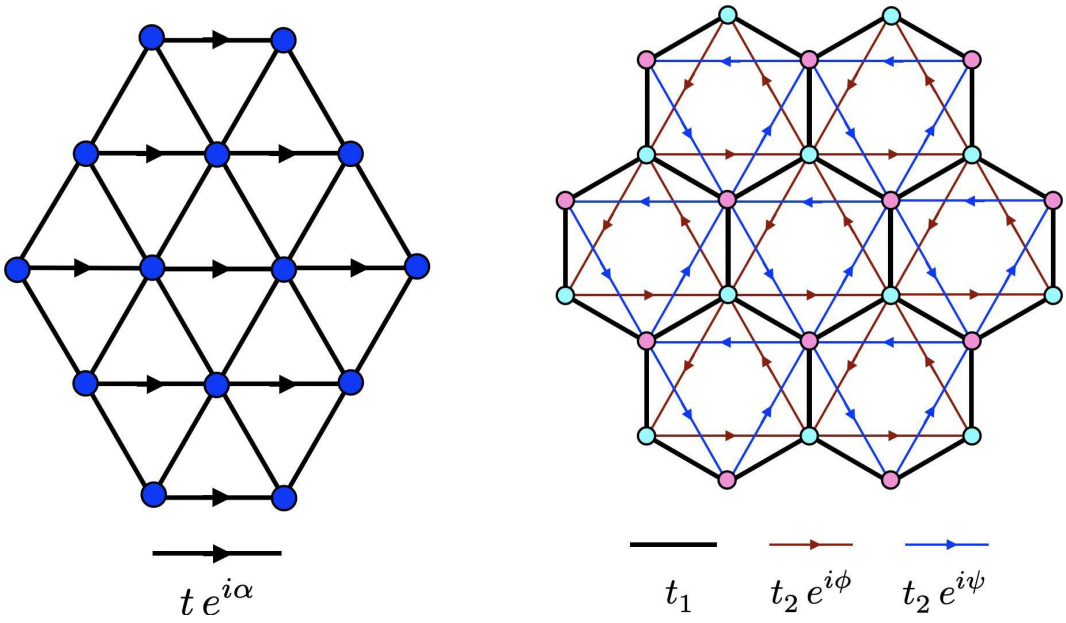


Figure 1.18: Two models with zero net flux per unit cell which still break time reversal symmetry. Left: The unit cell of the triangular lattice consists of two triangles.

closed loop of links on the lattice. Two counterexamples are shown in Fig. 1.18. The first example is that of the triangular lattice, where each structural unit cell is a rhombus consisting of two elementary triangular plaquettes. Consider now the situation where each horizontal link carries a U(1) phase α, *i.e.* $A_{\boldsymbol{r}\boldsymbol{r}'} = t\,e^{i\alpha}$, while the remaining links all have $A_{\boldsymbol{r}\boldsymbol{r}'} = 0$. Computing the U(1) flux by taking the directed sum counterclockwise over each triangle, we see that all the up triangles carry flux $\phi_\triangle = \alpha$, while all the down triangles carry flux $\phi_\triangledown = -\alpha \cong 2\pi - \alpha$. If, as before, we take the elementary direct lattice vectors to be $\boldsymbol{a}_{1,2} = a\big(\frac{1}{2}\hat{\boldsymbol{x}} \pm \frac{\sqrt{3}}{2}\hat{\boldsymbol{y}}\big)$ and write $\boldsymbol{k} = \sum_{j=1}^{2} \theta_j\,\boldsymbol{b}_j/2\pi$, with $\boldsymbol{a}_i \cdot \boldsymbol{b}_j = 2\pi\,\delta_{ij}$, then the tight binding Hamiltonian for this triangular lattice model is given by $H = \sum_{\boldsymbol{k}} E_{\boldsymbol{k}}\,a^\dagger_{\boldsymbol{k}}\,a_{\boldsymbol{k}}$, where

$$\begin{aligned} E_{\boldsymbol{k}} &= -2t\cos(\boldsymbol{k}\cdot\boldsymbol{a}_1 + \boldsymbol{k}\cdot\boldsymbol{a}_2 + \alpha) - 2t\cos(\boldsymbol{k}\cdot\boldsymbol{a}_1) - 2t\cos(\boldsymbol{k}\cdot\boldsymbol{a}_2) \\ &= -2t\cos(\theta_1 + \theta_2 + \alpha) - 2t\cos\theta_1 - 2t\cos\theta_2 \quad . \end{aligned} \tag{1.169}$$

A more interesting state of affairs is depicted in the right panel of Fig. 1.18, which is graphene augmented by nearest neighbor same-sublattice hopping terms, which is the celebrated Haldane honeycomb lattice model[27]. Inscribed in each hexagonal cell are one up-triangle of A site, depicted by blue dots in the figure, and one down-triangle of B sites, depicted as pink dots in the figure. Again we take $\boldsymbol{a}_{1,2} = a\big(\frac{1}{2}\hat{\boldsymbol{x}} \pm \frac{\sqrt{3}}{2}\hat{\boldsymbol{y}}\big)$ for the underlying A Bravais lattice, with the basis vectors $\boldsymbol{0}$ and $\boldsymbol{\delta}_1 = a\hat{\boldsymbol{y}}$. The nearest neighbor hoppings between A and B sites all are taken to have amplitude t_1, while the inscribed A and B same-sublattice hoppings are taken to be $t_2\,e^{i\phi}$

[27] F. D. M. Haldane, *Phys. Rev. Lett.* **61**, 2015 (1988).

and $t_2\,e^{i\psi}$, respectively, and taken in the counterclockwise direction around the inscribed $\triangle$ and $\triangledown$ paths. An on-site energy term $\pm m$, called the *Semenoff mass*, is added to the hopping terms. One then obtains the tight binding Hamiltonian

$$\begin{aligned} H = \sum_{\boldsymbol{k}} \Big\{ &- t_1 \left(1 + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right) a^\dagger_{\boldsymbol{k}}\, b_{\boldsymbol{k}} - t_1 \left(1 + e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right) b^\dagger_{\boldsymbol{k}}\, a_{\boldsymbol{k}} \\ &+ \Big[m - 2t_2\,\mathsf{Re}\left(e^{i\phi}\,e^{i\boldsymbol{k}\cdot(\boldsymbol{a}_1+\boldsymbol{a}_2)} + e^{i\phi}\,e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\phi}\,e^{-i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right)\Big] a^\dagger_{\boldsymbol{k}}\, a_{\boldsymbol{k}} \\ &\Big[-m - 2t_2\,\mathsf{Re}\left(e^{i\psi}\,e^{-i\boldsymbol{k}\cdot(\boldsymbol{a}_1+\boldsymbol{a}_2)} + e^{i\psi}\,e^{i\boldsymbol{k}\cdot\boldsymbol{a}_1} + e^{i\psi}\,e^{i\boldsymbol{k}\cdot\boldsymbol{a}_2}\right)\Big] b^\dagger_{\boldsymbol{k}}\, b_{\boldsymbol{k}} \Big\} \\ = \sum_{\boldsymbol{k}} &\begin{pmatrix} a^\dagger_{\boldsymbol{k}} & b^\dagger_{\boldsymbol{k}} \end{pmatrix} \begin{pmatrix} H_{\mathsf{AA}}(\boldsymbol{k}) & H_{\mathsf{AB}}(\boldsymbol{k}) \\ H_{\mathsf{BA}}(\boldsymbol{k}) & H_{\mathsf{BB}}(\boldsymbol{k}) \end{pmatrix} \begin{pmatrix} a_{\boldsymbol{k}} \\ b_{\boldsymbol{k}} \end{pmatrix} \quad , \end{aligned} \tag{1.170}$$

where

$$\begin{aligned} H_{\mathsf{AA}}(\boldsymbol{k}) &= m - 2t_2\big[\cos\theta_1 + \cos\theta_2 + \cos(\theta_1+\theta_2)\big]\cos\phi - 2t_2\big[\sin\theta_1 + \sin\theta_2 - \sin(\theta_1+\theta_2)\big]\sin\phi \\ H_{\mathsf{BB}}(\boldsymbol{k}) &= -m - 2t_2\big[\cos\theta_1 + \cos\theta_2 + \cos(\theta_1+\theta_2)\big]\cos\psi + 2t_2\big[\sin\theta_1 + \sin\theta_2 - \sin(\theta_1+\theta_2)\big]\sin\psi \\ H_{\mathsf{AB}}(\boldsymbol{k}) &= H^*_{\mathsf{BA}}(\boldsymbol{k}) = -t_1\big(1 + e^{i\theta_1} + e^{-i\theta_2}\big) \quad . \end{aligned} \tag{1.171}$$

In the Haldane model, $\psi = \phi$, in which case we may write $H(\boldsymbol{k})$ in terms of Pauli matrices, as

$$\begin{aligned} H(\boldsymbol{k}) = &-2t_2\big[\cos\theta_1 + \cos\theta_2 + \cos(\theta_1+\theta_2)\big]\cos\phi - t_1\big(1 + \cos\theta_1 + \cos\theta_2\big)\,\sigma^x \\ &+ t_1\big(\sin\theta_1 - \sin\theta_2\big)\,\sigma^y + \Big(m - 2t_2\big[\sin\theta_1 + \sin\theta_2 - \sin(\theta_1+\theta_2)\big]\sin\phi\Big)\,\sigma^z \quad . \end{aligned} \tag{1.172}$$

What makes the Haldane model so interesting is that its band structure is *topological* over a range of the dimensionless parameters m/t_2 and ϕ. (Without loss of generality, we may set $t_1 \equiv 1$.) More on this below!

1.6.4 General flux configuration on the square lattice

More generally, consider a magnetic unit cell formed by an $M \times N$ block of structural unit cells, as depicted in Fig. 1.19. Each structural cell p is labeled by the indices (m, n), where the Bravais lattice site its lower left corner is $ma\hat{\boldsymbol{x}} + na\hat{\boldsymbol{y}}$. To assign the lattice gauge fields, do the following. For $\boldsymbol{r} = ma\hat{\boldsymbol{x}} + na\hat{\boldsymbol{y}}$ and $\boldsymbol{r}' = ma\hat{\boldsymbol{x}} + (n+1)a\hat{\boldsymbol{y}}$ with $n < N$, let $A_{\boldsymbol{r}\boldsymbol{r}'} = \sum_{i=1}^{m-1}\phi_{i,n}$. For the $n = N$, we include the Bloch phase θ_2, so that $A_{\boldsymbol{r}\boldsymbol{r}'} = \theta_2 + \sum_{i=1}^{m-1}\phi_{i,n}$, also noting that $(m, N+1) \cong (m, 1)$. This sets $A_{\boldsymbol{r}\boldsymbol{r}'}$ for all vertical (y-directed) links. The only horizontal links for which $A_{\boldsymbol{r}\boldsymbol{r}'}$ are nonzero are those with $\boldsymbol{r} = Ma\hat{\boldsymbol{x}} + na\hat{\boldsymbol{y}}$ and $\boldsymbol{r}' = a\hat{\boldsymbol{x}} + na\hat{\boldsymbol{y}}$; note $(M+1, n) \cong (1, n)$. Then $A_{\boldsymbol{r}\boldsymbol{r}'} = \theta_1 - \sum_{i=1}^{M}\sum_{j=1}^{n}\phi_{i,j}$. One can check that this prescription yields the desired flux configuration, as well as the two Bloch phases.

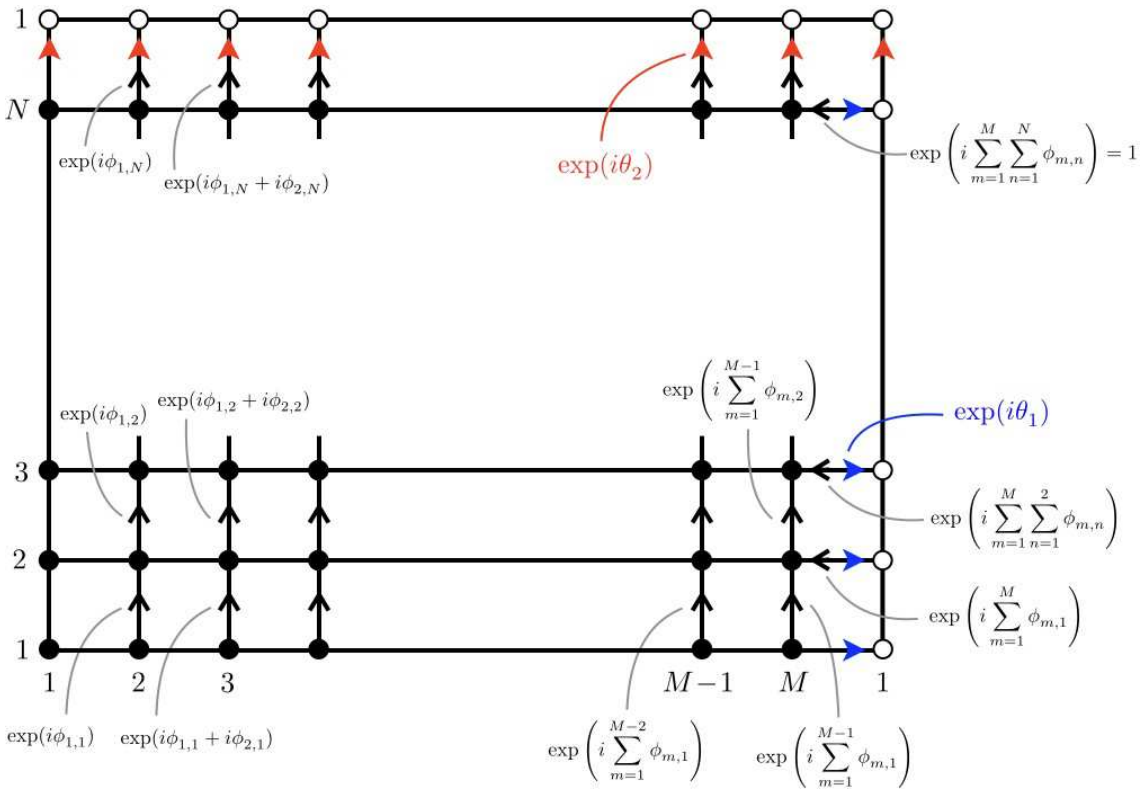


Figure 1.19: Lattice gauge field configuration for a general $M \times N$ rectangular lattice with flux $\phi_{m,n}$ in unit cell (m, n) and Bloch phases (θ_1, θ_2).

1.7 Berry Phases, Fiber Bundles, Chern Numbers, and TKNN

1.7.1 The adiabatic theorem and Berry's phase

Consider a Hamiltonian $H(\boldsymbol{\lambda})$ dependent on a set of parameters $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_K\}$, and let $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle$ satisfy the time-independent Schrödinger equation,

$$H(\boldsymbol{\lambda})\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle = E_n(\boldsymbol{\lambda})\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle \quad . \tag{1.173}$$

Now let $\boldsymbol{\lambda}(t)$ be continuously time-dependent, and consider the time-*dependent* Schrödinger equation

$$i\hbar \frac{d}{dt}\,|\,\Psi(t)\,\rangle = H\big(\boldsymbol{\lambda}(t)\big)\,|\,\Psi(t)\,\rangle \quad . \tag{1.174}$$

The adiabatic theorem states that if $\boldsymbol{\lambda}(t)$ evolves extremely slowly, then each solution $|\,\Psi_n(t)\,\rangle$ is proportional to $|\,\varphi_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle$, with

$$|\,\Psi_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle = \exp\big(i\gamma_n(t)\big)\,\exp\Big(-\frac{i}{\hbar}\int^t dt'\,E_n\big(\boldsymbol{\lambda}(t')\big)\Big)\,|\,\varphi_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle \quad , \tag{1.175}$$

with corrections which vanish in the limit $|\dot{\boldsymbol{\lambda}}|/|\boldsymbol{\lambda}| \to 0$. Taking the time derivative and then the overlap with the bra vector $\langle\, \varphi_n\big(\boldsymbol{\lambda}(t)\big)\,|$, one obtains the result

$$\frac{d\gamma_n(t)}{dt} = i\,\langle\, \varphi_n\big(\boldsymbol{\lambda}(t)\big)\,|\,\frac{d}{dt}\,|\,\varphi_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle = \boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda})\cdot\frac{d\boldsymbol{\lambda}}{dt} \equiv \mathcal{A}_n(t) \quad , \tag{1.176}$$

where

$$\mathcal{A}_n^\mu(\boldsymbol{\lambda}) = i\,\langle\, \varphi_n(\boldsymbol{\lambda})\,|\,\frac{\partial}{\partial\lambda_\mu}\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle \quad . \tag{1.177}$$

Note that $\mathcal{A}_n^\mu(\boldsymbol{\lambda})$ is real. In particular, if $\boldsymbol{\lambda}(t)$ traverses a closed loop $\mathcal{C}$ with infinitesimal speed, then the wavefunction $|\,\Psi_n(t)\,\rangle$ will accrue a *geometric phase* $\gamma_n(\mathcal{C})$, given by

$$\gamma_n(\mathcal{C}) = \oint\limits_{\mathcal{C}} d\boldsymbol{\lambda}\cdot\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda}) \quad , \tag{1.178}$$

called *Berry's phase*[28].

In the adiabatic limit, the dynamical phase $\hbar^{-1}\int^t dt'\,E_n\big(\boldsymbol{\lambda}(t')\big)$ becomes very large if $E_n \neq 0$, because the path $\boldsymbol{\lambda}(t)$ is traversed very slowly. We may remove this dynamical phase by defining the Hamiltonian

$$\widetilde{H}_n(\boldsymbol{\lambda}) \equiv H(\boldsymbol{\lambda}) - E_n(\boldsymbol{\lambda}) \quad . \tag{1.179}$$

We define $|\,\widetilde{\Psi}_n(t)\,\rangle$ as the solution to the Schrödinger equation

$$i\hbar\,\frac{d}{dt}\,|\,\widetilde{\Psi}_n(t)\,\rangle = \widetilde{H}_n\big(\boldsymbol{\lambda}(t)\big)\,|\,\widetilde{\Psi}_n(t)\,\rangle \tag{1.180}$$

in the adiabatic limit. The adiabatic wavefunctions $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle$ are the same as before, but now satisfy the zero energy condition $\widetilde{H}_n(\boldsymbol{\lambda})\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle = 0$. Clearly $|\,\widetilde{\Psi}_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle = \exp\big(i\gamma_n(t)\big)\,|\,\varphi_n\big(\boldsymbol{\lambda}(t)\big)\,\rangle$ and the dynamical phase has been removed. However, note that the geometrical phase γ_n does not depend on the elapsed time, but only on the path traversed, *viz.*

$$\gamma_n = \gamma_n(\boldsymbol{\lambda}) = \int\limits_{\boldsymbol{\lambda}_0}^{\boldsymbol{\lambda}} d\boldsymbol{\lambda}'\cdot\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda}') \quad , \tag{1.181}$$

where $\boldsymbol{\lambda}_0 = \boldsymbol{\lambda}(0)$, and where the integral is taken along the path in traversed by $\boldsymbol{\lambda}$.

1.7.2 Connection and curvature

The mathematical structure underlying this discussion is that of the *Hermitian line bundle*, the ingredients of which are (i) a *base space* $\mathcal{M}$ which is a topological manifold; this is the parameter

[28] See M. V. Berry, *Proc. Roy. Soc. A* **392**, 45 (1984).

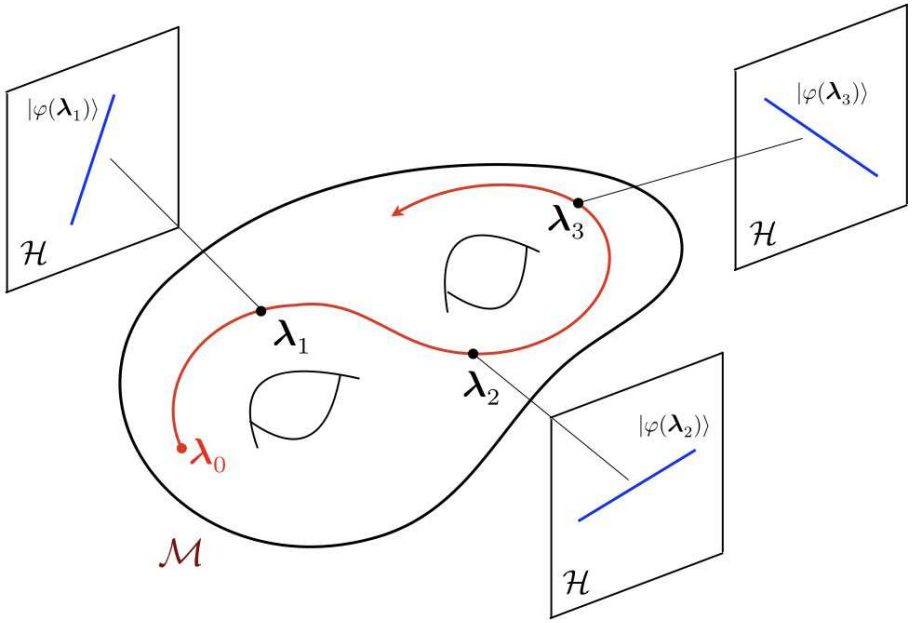


Figure 1.20: A Hermitian line bundle consists of a base space $\mathcal{M}$ and a fiber $|\,\varphi(\boldsymbol{\lambda})\,\rangle$ which twists as the point $\boldsymbol{\lambda}$ moves around the base space.

space for $\boldsymbol{\lambda}$, and (ii) to each point $\boldsymbol{\lambda} \in \mathcal{M}$ is associated a *fiber* which is the adiabatic wavefunction $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle \in \mathcal{H}$, which is a complex one-dimensional subspace of some Hilbert space $\mathcal{H}$. As $\boldsymbol{\lambda}$ moves around the base space $\mathcal{M}$, the fiber twists around. The adiabatic theorem furnishes us with a way of defining *parallel transport* of $|\,\widetilde{\Psi}_n(\boldsymbol{\lambda})\,\rangle$ along the curve $\mathcal{C}$[29]. The object $\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda})$ is the *connection* and the geometric phase $\gamma(\mathcal{C})$ is the *holonomy* of the connection[30]. As a holonomy, $\gamma(\mathcal{C})$ depends only on the curve $\mathcal{C}$ and not on where along the curve one starts.

The *curvature tensor* for the bundle is given by

$$\begin{aligned}\Omega_n^{\mu\nu}(\boldsymbol{\lambda}) &= \frac{\partial \mathcal{A}_n^\nu}{\partial \lambda_\mu} - \frac{\partial \mathcal{A}_n^\mu}{\partial \lambda_\nu} \\ &= i\left\langle \frac{\partial \varphi_n}{\partial \lambda_\mu} \,\middle|\, \frac{\partial \varphi_n}{\partial \lambda_\nu} \right\rangle - i\left\langle \frac{\partial \varphi_n}{\partial \lambda_\nu} \,\middle|\, \frac{\partial \varphi_n}{\partial \lambda_\mu} \right\rangle \quad .\end{aligned} \tag{1.182}$$

Using completeness of the $|\,\varphi_n\,\rangle$ basis, we may write the curvature tensor as

$$\Omega_n^{\mu\nu}(\boldsymbol{\lambda}) = i \sum_l{}' \frac{\left\langle \varphi_n \,\middle|\, \frac{\partial H}{\partial \lambda_\mu} \,\middle|\, \varphi_l \right\rangle \left\langle \varphi_l \,\middle|\, \frac{\partial H}{\partial \lambda_\nu} \,\middle|\, \varphi_n \right\rangle - (\mu \leftrightarrow \nu)}{(E_n - E_l)^2} \quad , \tag{1.183}$$

[29]Note that $|\,\Psi_n(t)\,\rangle$, which depends explicitly on elapsed time and not solely on the position $\boldsymbol{\lambda}$ along its trajectory, can not be said to be parallel transported along any curve.

[30]See B. Simon, *Phys. Rev. Lett.* **51**, 2167 (1983).

where the prime on the sum indicates that the term $l = n$ is to be excluded. We see that in this formulation the curvature tensor is actually independent of any phase convention for the adiabatic wavefunctions $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle$. So long as the adiabatic eigenstate $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle$ remains nondegenerate, the denominator in Eqn. 1.183 remains nonzero, hence the curvature tensor $\Omega(\boldsymbol{\lambda})$ is nonsingular. The same cannot be said about the connection $\boldsymbol{\mathcal{A}}(\boldsymbol{\lambda})$, however, because it is *gauge-covariant*. This means that if we multiply the adiabatic wavefunctions by phases, with $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle \to \exp(if_n(\boldsymbol{\lambda}))\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle$, the connection changes accordingly, with

$$\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda}) \to \boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda}) - \frac{\partial f_n(\boldsymbol{\lambda})}{\partial \boldsymbol{\lambda}} \quad . \tag{1.184}$$

How can we fix a gauge in order to give unambiguous meaning to $\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda})$? One way might be to demand that the adiabatic wavefunction amplitude be real and positive at some particular point in space $\boldsymbol{r}_0$, *i.e.* $\langle\,\boldsymbol{r}_0\,|\,\varphi_n(\boldsymbol{\lambda})\,\rangle \in \mathbb{R}_+$ for all $\boldsymbol{\lambda} \in \mathcal{M}$. For lattice-based models, where the adiabatic wavefunction is a vector of amplitudes for each orbital and each site within the appropriate unit cell, we could similarly demand that one of these amplitudes be real and positive. *This prescription fails if there exists a value of $\boldsymbol{\lambda}$ for which this wavefunction amplitude vanishes.*

As we are about to discover, the integral of the curvature over a two-dimensional base space is a topological invariant, meaning that it remains fixed (and indeed quantized) under continuous deformations of the Hamiltonian $H(\boldsymbol{\lambda})$. Using Stokes' theorem, we can turn an area integral of the curvature into line integrals of the connection. However, having chosen a particular gauge for the adiabatic wavefunctions, it may be that the connection is singular at certain points. Therefore the line integrals cannot be completely collapsed, and we obtain the result

$$\int\limits_{\mathcal{M}} d^2\lambda\; \Omega_n^{12}(\boldsymbol{\lambda}) = -\sum_i \oint\limits_{\mathcal{C}_i} d\boldsymbol{\lambda} \cdot \boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda}) \quad , \tag{1.185}$$

where the loop $\mathcal{C}_i$ encloses the $i^{\rm th}$ singularity $\boldsymbol{\lambda}_i$ of the connection in a counterclockwise manner[31]. This is the generalization to Hermitian line bundles of the index formula in Eqn. 1.288 for the Gauss-Bonnet theorem. Quantization follows by writing $|\,\varphi_n(\boldsymbol{\lambda})\,\rangle = e^{iq_i\zeta(\boldsymbol{\lambda}-\boldsymbol{\lambda}_i)}\,|\,\widetilde{\varphi}_n(\boldsymbol{\lambda})\,\rangle$ in the vicinity of $\boldsymbol{\lambda} = \boldsymbol{\lambda}_i$, where q_i is an integer and

$$\zeta(\boldsymbol{\lambda} - \boldsymbol{\lambda}_i) = \tan^{-1}\left(\frac{\lambda_2 - \lambda_{i,2}}{\lambda_1 - \lambda_{i,1}}\right) \quad . \tag{1.186}$$

The integers q_i are chosen to 'unwind' the singularity at each $\boldsymbol{\lambda}_i$, so as to make the gauge transformed connection $\widetilde{\boldsymbol{\mathcal{A}}}_n^{\mu}(\boldsymbol{\lambda}) \equiv i\,\langle\,\widetilde{\varphi}_n(\boldsymbol{\lambda})\,|\,\boldsymbol{\nabla}_{\boldsymbol{\lambda}}\,|\,\widetilde{\varphi}_n(\boldsymbol{\lambda})\,\rangle$ nonsingular[32]. We then obtain

$$C_n \equiv \frac{1}{2\pi}\int\limits_{\mathcal{M}} d^2\lambda\; \Omega_n^{12}(\boldsymbol{\lambda}) = \sum_i q_i \quad . \tag{1.187}$$

[31] We must assume that the base space $\mathcal{M}$ is orientable.

[32] Note that we have employed a singular gauge transformation, which is necessary to do the desired unwinding. Also note that the integers q_i should also carry a band index n, which we have suppressed here for notational simplicity.

Thus $C_n \in \mathbb{Z}$ is the *Chern number* of the Hermitian line bundle corresponding to the adiabatic wavefunction $|\,\varphi_n\,\rangle$.

The simplest nontrivial example is that of a spin-$\frac{1}{2}$ object in a magnetic field $\boldsymbol{B}(t)$, with

$$H(t) = g\mu_{\rm B}\boldsymbol{B}\cdot\boldsymbol{\sigma} = g\mu_{\rm B}B\begin{pmatrix}\cos\theta & \sin\theta\exp(-i\phi)\\ \sin\theta\exp(i\phi) & \cos\theta\end{pmatrix} \quad , \tag{1.188}$$

where $\boldsymbol{B} = B\,\hat{\boldsymbol{n}}$ is the adiabatic parameter which varies extremely slowly in time. The adiabatic wavefunctions are

$$|\,\varphi_+(\theta,\phi)\,\rangle = \begin{pmatrix}u\\ v\end{pmatrix} \qquad , \qquad |\,\varphi_-(\theta,\phi)\,\rangle = \begin{pmatrix}-\bar{v}\\ \bar{u}\end{pmatrix} \quad , \tag{1.189}$$

where $\hat{\boldsymbol{n}} = (\sin\theta\cos\phi\,,\,\sin\theta\sin\phi\,,\,\cos\theta)$, $u = \cos(\frac{1}{2}\theta)$, and $v = \sin(\frac{1}{2}\theta)\exp(i\phi)$. The energy eigenvalues are $E_\pm = \pm g\mu_{\rm B}B$. We now compute the connections,

$$\begin{aligned}\mathcal{A}_+ &= i\,\langle\,\varphi_+\,|\,\frac{d}{dt}\,|\,\varphi_+\,\rangle = i\big(\bar{u}\dot{u} + \bar{v}\dot{v}\big) = -\tfrac{1}{2}(1-\cos\theta)\,\dot{\phi} = -\tfrac{1}{2}\,\dot{\omega}\\ \mathcal{A}_- &= i\,\langle\,\varphi_-\,|\,\frac{d}{dt}\,|\,\varphi_-\,\rangle = i\big(u\dot{\bar{u}} + v\dot{\bar{v}}\big) = +\tfrac{1}{2}(1-\cos\theta)\,\dot{\phi} = +\tfrac{1}{2}\,\dot{\omega} \quad ,\end{aligned} \tag{1.190}$$

where $\dot{\omega}$ is the differential solid angle subtended by the path $\hat{\boldsymbol{n}}(t)$. Thus, $\gamma_\pm(\mathcal{C}) = \mp\frac{1}{2}\,\omega_{\mathcal{C}}$ is $\mp$ half the solid angle subtended by the path $\hat{\boldsymbol{n}}_{\mathcal{C}}(t)$ on the Bloch sphere. We may now read off the components $\mathcal{A}^\theta_\pm = 0$ and $\mathcal{A}^\phi_\pm = \mp\frac{1}{2}(1-\cos\theta)$ and invoke Eqn. 1.182 to compute the curvature,

$$\Omega^{\theta\phi}_\pm(\theta,\phi) = \mp\tfrac{1}{2}\sin\theta \quad . \tag{1.191}$$

Note then that the integral of the curvature over the entire sphere is given by

$$\int\limits_0^{2\pi}\!d\phi\int\limits_0^{\pi}\!d\theta\;\Omega^{\theta\phi}_\pm(\theta,\phi) = 2\pi\,C_\pm \quad , \tag{1.192}$$

where $C_\pm = \mp 1$ is the Chern number. Equivalently, note that both connections are singular at $\theta = \pi$, where the azimuthal angle is ill-defined. This singularity can be gentled through an appropriate singular gauge transformation $|\,\varphi_\pm\,\rangle = e^{\pm i\phi}\,|\,\widetilde{\varphi}_\pm\,\rangle = e^{\mp i\zeta}\,|\,\widetilde{\varphi}_\pm\,\rangle$, where ζ is defined to be the angle which increases as one winds counterclockwise around the *south pole*, hence $\zeta = -\phi$. This corresponds to $q_\pm = \mp 1$ in our earlier notation, hence again $C_\pm = \mp 1$.

As we saw above, this is a general result: when the base space $\mathcal{M}$ is two dimensional: the integral of the curvature over $\mathcal{M}$ is 2π times an integer. This result calls to memory the famous Gauss-Bonnet theorem (see §1.10 below for more), which says that the integral of the Gaussian curvature K over a two-dimensional manifold $\mathcal{M}$ is

$$\int\limits_{\mathcal{M}} dS\;K = 4\pi(1-g) \quad , \tag{1.193}$$

where g is the *genus* (number of holes) in the manifold $\mathcal{M}$. In the Gauss-Bonnet case, the bundle construction is known as the *tangent bundle* of $\mathcal{M}$, and the corresponding connection is determined by the Riemannian metric one places on $\mathcal{M}$. However, *independent of the metric*, the integral of K is determined solely by the global topology of $\mathcal{M}$, *i.e.* by its genus. Thus, in three-dimensional space, a sphere inherits a metric from its embedding in $\mathbb{R}^3$. If you distort the sphere by denting it, locally its curvature K will change, being the product of the principal radii of curvature at any given point. But the integral of K over the surface will remain fixed at 4π. Just as the genus g of a Riemann surface is unaffected by simple deformations but can change if one does violence to it, such as puncturing and resewing it[33], so is the Chern number invariant under deformations of the underlying Hamiltonian, provided one does not induce a level crossing of the adiabatic eigenstate $|\,\varphi_n\,\rangle$ with a neighboring one. Also, note that if the connection $\boldsymbol{\mathcal{A}}_n(\boldsymbol{\lambda})$ can be defined globally on $\mathcal{M}$, *i.e.* with no singularities, then $C_n = 0$.

1.7.3 Two-band models

For the two band ($S = \frac{1}{2}$) system with Hamiltonian $H = g\mu_{\rm B} B\,\hat{\boldsymbol{n}}(\boldsymbol{\lambda})\cdot\boldsymbol{\sigma}$, one can verify that we may also write the Chern numbers as[34]

$$C_\pm = \pm\frac{1}{4\pi}\int\limits_{\mathcal{M}} d^2\lambda\;\hat{\boldsymbol{n}}\cdot\frac{\partial\hat{\boldsymbol{n}}}{\partial\lambda_1}\times\frac{\partial\hat{\boldsymbol{n}}}{\partial\lambda_2}\quad . \tag{1.194}$$

In this formulation, the Chern number has the interpretation of a *Pontrjagin number*, which is a topological index classifying real vector bundles (more in §1.10.2 below). Thus, for a tight binding model on a bipartite lattice, the most general Hamiltonian may be written

$$H(\boldsymbol{k}) = d_0(\boldsymbol{k}) + \boldsymbol{d}(\boldsymbol{k})\cdot\boldsymbol{\sigma}\quad , \tag{1.195}$$

where $\boldsymbol{k}$ is the wavevector and where each $d^\mu(\boldsymbol{k})$ is periodic under translations of $\boldsymbol{k}$ by any reciprocal lattice vector $\boldsymbol{G}$. In this case $\lambda_{1,2} = k_{x,y}$ are the components of $\boldsymbol{k}$, and $\mathcal{M} = \mathbb{T}^2$ is the Brillouin zone torus. Note that the sum of the Chern numbers for each of the $+$ and $-$ bands is zero. As we shall see below with the TKNN problem, for a larger spin generalization, *i.e.* when the magnetic unit cell contains more than two basis elements, the sum $\sum_a C_a$ of the Chern numbers over all bands also vanishes. Consider the two band model with

$$H(\boldsymbol{\theta}) = \begin{pmatrix} m - 2t\cos\theta_1 - 2t\cos\theta_2 & \Delta\,(\sin\theta_1 - i\sin\theta_2) \\ \Delta\,(\sin\theta_1 + i\sin\theta_2) & -m + 2t\cos\theta_1 + 2t\cos\theta_2 \end{pmatrix}\quad . \tag{1.196}$$

As before, $\theta_\mu = \boldsymbol{k}\cdot\boldsymbol{a}_\mu$. Note $H(\boldsymbol{\theta}) = \boldsymbol{d}(\boldsymbol{\theta})\cdot\boldsymbol{\sigma}$ with

$$\begin{aligned}\boldsymbol{d}(\boldsymbol{\theta}) &= \big(\Delta\sin\theta_1\,,\,\Delta\sin\theta_2\,,\,m - 2t\cos\theta_1 - 2t\cos\theta_2\big) \\ &\equiv |\boldsymbol{d}|\,\big(\sin\vartheta\cos\phi\,,\,\sin\vartheta\sin\phi\,,\,\cos\vartheta\big)\quad .\end{aligned} \tag{1.197}$$

[33] M. Gilbert's two commandments of topology: (I) Thou shalt not cut. (II) Thou shalt not glue.

[34] The dependence of the magnitude $B = |\boldsymbol{B}|$ on $\boldsymbol{\lambda}$ is irrelevant to the calculation of the Chern numbers. The equivalence $\hat{\boldsymbol{n}} = z^\dagger\boldsymbol{\sigma}z$ for $z = \begin{pmatrix} u \\ v\end{pmatrix}$ is known as the *first Hopf map* from CP^1 to $\mathbb{S}^2$.

Note the adiabatic parameters here are θ_1 and θ_2, upon which ϑ and ϕ are parametrically dependent. Does $\boldsymbol{d}(\boldsymbol{\theta})$ wind around the Brillouin zone torus, yielding a nonzero Chern number?

First, you might be wondering, where does this model come from? Actually, it is the Hamiltonian for a $p_x + ip_y$ superconductor, but we can back out of $H(\boldsymbol{\theta})$ a square lattice insulator model involving two orbitals a and b which live on top of each other at each site, and are not spatially separated[35]. The parameter m reflects the difference in the local energies of the two orbitals. The nearest neighbor hopping integrals between like orbitals are $t_{aa} = t$ and $t_{bb} = -t$, but $t_{ab}(\pm\boldsymbol{a}_1) = \pm\frac{i}{2}\,\Delta$ and $t_{ab}(\pm\boldsymbol{a}_2) = \pm\frac{1}{2}\Delta$, with $t_{ba}(-\boldsymbol{a}_{1,2}) = t^*_{ab}(+\boldsymbol{a}_{1,2})$ due to hermiticity.

The energy eigenvalues are

$$E_\pm(\boldsymbol{\theta}) = \pm\sqrt{\Delta^2\sin^2\theta_1 + \Delta^2\sin^2\theta_2 + (m - 2t\cos\theta_1 - \cos\theta_2)^2} \quad . \tag{1.198}$$

The Wigner-von Neumann theorem says that degeneracy for complex Hamiltonians like ours has codimension three, meaning one must fine tune three parameters in order to get a degeneracy. The reason is that for $H = \boldsymbol{d}\cdot\sigma$ describing two nearby levels, the gap is $2|\boldsymbol{d}|$, thus in order for the gap to vanish we must require three conditions: $d_x = d_y = d_z = 0$. For the real case where $d_y = 0$ is fixed, we only require two conditions, *i.e.* $d_x = d_z = 0$. For our model, the gap collapse requires

$$\begin{aligned} \Delta\,\sin\theta_1 &= 0 \\ \Delta\,\sin\theta_2 &= 0 \\ m - 2t\cos\theta_1 - 2t\cos\theta_2 &= 0 \quad . \end{aligned} \tag{1.199}$$

Thus, degeneracies occur at $(\theta_1,\theta_2) = (0,0)$ when $m = 4t$, at (π,π) when $m = -4t$, and at $(0,\pi)$ and $(\pi,0)$ when $m = 0$. It is clear that for $|m| > 4t$ both Chern numbers must be zero. This is because for $m > 4t$ we have $d_z(\theta_1,\theta_2) > 0$ for all values of the Bloch phases, while for $m < -4t$ we have $d_z(\theta_1,\theta_2) < 0$. Thus in neither case can the $\boldsymbol{d}$ vector wind around the Bloch sphere, and the Pontrjagin/Chern indices accordingly vanish for both bands.

Now consider the case $m \in [0, 4t]$. Recall that the eigenfunctions are given by

$$|\,\varphi_+\,\rangle = \begin{pmatrix} \cos(\frac{1}{2}\vartheta) \\ \sin(\frac{1}{2}\vartheta)\,e^{i\chi} \end{pmatrix} \qquad , \qquad |\,\varphi_-\,\rangle = \begin{pmatrix} -\sin(\frac{1}{2}\vartheta)\,e^{-i\chi} \\ \cos(\frac{1}{2}\vartheta) \end{pmatrix} \quad , \tag{1.200}$$

with eigenvalues $\pm|\boldsymbol{d}|$. The singularity in both $|\,\varphi_\pm(\theta_1,\theta_2)\,\rangle$ occurs at $\vartheta = \pi$. Recall that $\boldsymbol{d} \equiv |\boldsymbol{d}|\big(\sin\vartheta\cos\chi\,,\,\sin\vartheta\sin\chi\,,\,\cos\vartheta\big)$, which entails $\boldsymbol{d} = (0\,,\,0\,,\,-|\boldsymbol{d}|)$, *i.e.* $d_x = d_y = 0$ and $d_z < 0$. This only occurs for $(\theta_1,\theta_2) = (0,0)$. All we need to do to compute the Chern numbers is to identify the singularity in $\zeta(\theta_1,\theta_2)$ about this point, *i.e.* does $\zeta = -\chi$ wind clockwise or counterclockwise, in which case $C_+ = -1$ or $C_+ = +1$, respectively. Treating $\theta_{1,2}$ as very small, one easily obtains $\zeta = -\tan^{-1}(\theta_2/\theta_1)$, which is to say clockwise winding, hence $C_\pm = \mp 1$. *Exercise*: Find $C_\pm$ for $m \in [-4t, 0]$.

[35] In this model they are both s-orbitals, which is unphysical.

Haldane honeycomb model

In §1.6.3, we met Haldane's famous honeycomb lattice model, $H(\boldsymbol{\theta}) = d_0(\boldsymbol{\theta}) + \boldsymbol{d}(\boldsymbol{\theta}) \cdot \boldsymbol{\sigma}$, with

$$\begin{aligned} d_0(\boldsymbol{\theta}) &= -2t_2\big[\cos\theta_1 + \cos\theta_2 + \cos(\theta_1+\theta_2)\big]\cos\phi \\ d_x(\boldsymbol{\theta}) &= -t_1\big(1 + \cos\theta_1 + \cos\theta_2\big) \\ d_y(\boldsymbol{\theta}) &= t_1\big(\sin\theta_1 - \sin\theta_2\big) \\ d_z(\boldsymbol{\theta}) &= m - 2t_2\big[\sin\theta_1 + \sin\theta_2 - \sin(\theta_1+\theta_2)\big]\sin\phi \quad . \end{aligned} \tag{1.201}$$

The energy eigenvalues are $E_\pm(\boldsymbol{\theta}) = d_0(\boldsymbol{\theta}) \pm |\boldsymbol{d}(\boldsymbol{\theta})|$. Now is quite easy to demonstrate that $\big|\sin\theta_1 + \sin\theta_2 - \sin(\theta_1+\theta_2)\big| \le \frac{3}{2}\sqrt{3}$, and therefore that the $\boldsymbol{d}(\boldsymbol{\theta})$ cannot wind if $|m| > 3\sqrt{3}\,t_2\,|\sin\phi|$ and $C_\pm = 0$. As above, we set $\boldsymbol{d} \equiv |\boldsymbol{d}|\big(\sin\vartheta\cos\chi\,,\,\sin\vartheta\sin\chi\,,\,\cos\vartheta\big)$, and the singularity in both wavefunctions occurs at $\vartheta = \pi$, which requires $d_x(\boldsymbol{\theta}) = d_y(\boldsymbol{\theta}) = 0$ and $d_z(\boldsymbol{\theta}) < 0$. This in turn requires $\theta_1 = \theta_2 = \frac{2}{3}\pi s$ where $s = \mathsf{sgn}(\sin\phi)$. We now write $\theta_j = \frac{2}{3}\,\pi s + \delta_j$ and find

$$\tan\chi = \frac{d_y(\boldsymbol{\theta})}{d_x(\boldsymbol{\theta})} = \frac{\delta_2 - \delta_1}{\delta_1 + \delta_2}\,\mathsf{sgn}(\sin\phi) = \tan(\alpha - \tfrac{\pi}{4})\,\mathsf{sgn}(\sin\phi) \quad , \tag{1.202}$$

where $\boldsymbol{\delta} \equiv |\boldsymbol{\delta}|\,(\cos\alpha, \sin\alpha)$. Thus the winding of $\zeta = -\chi$ is in the sense of that of α if $s < 0$ and opposite if $s > 0$, and we conclude $C_\pm = \mp\mathsf{sgn}(\sin\phi)$. The topological phase diagram for the Haldane honeycomb lattice model is shown in Fig. 1.21. The phase space is a cylinder in the dimensionless parameters $\phi \in [-\pi, \pi]$ and $m/t_2 \in \mathbb{R}$. Regions are labeled by the Chern numbers $C_\pm$ of the two energy bands.

Note on broken symmetries

In §1.4.3 we derived the long wavelength Hamiltonian for the $\boldsymbol{K}$ and $\boldsymbol{K}' = -\boldsymbol{K}$ valleys of graphene. If we adopt a pseudospin convention for the valleys, with Pauli matrices $\boldsymbol{\tau}$, and where $\tau^z = \pm 1$ corresponds to the $\pm\boldsymbol{K}$ valley, we may in one stroke write the long wavelength graphene Hamiltonian as

$$H_0 = \tfrac{\sqrt{3}}{2}\,ta\big(q_x\,\sigma^x\,\tau^z + q_y\,\sigma^y\big) \quad . \tag{1.203}$$

This Hamiltonian is symmetric under the operations of parity ($\mathcal{P}$) and time-reversal ($\mathcal{T}$). Under $\mathcal{P}$, we switch valleys, switch sublattices, and send $q_x \to -q_x$. Under $\mathcal{T}$, we switch valleys and send $\boldsymbol{q} \to -\boldsymbol{q}$. It is also important to remember that $\mathcal{T}$ is antiunitary, and includes the complex conjugation operator $\check{K}$. The matrix parts of these operators, *i.e.* other than their actions on the components of $\boldsymbol{q}$, are given by

$$\mathcal{P} = \sigma^y\,\tau^z \qquad , \qquad \mathcal{T} = i\tau^y\check{K} \quad . \tag{1.204}$$

Note that $\mathcal{T}^2 = -1$ and $\mathcal{T}^{-1} = -\mathcal{T} = \check{K}\tau^y\,(-i)$. Of course $\mathcal{P}^2 = 1$ and thus $\mathcal{P}^{-1} = \mathcal{P}$. One can now check explicitly that $\mathcal{P}H_0\mathcal{P}^{-1} = \mathcal{T}H_0\mathcal{T}^{-1} = H_0$.

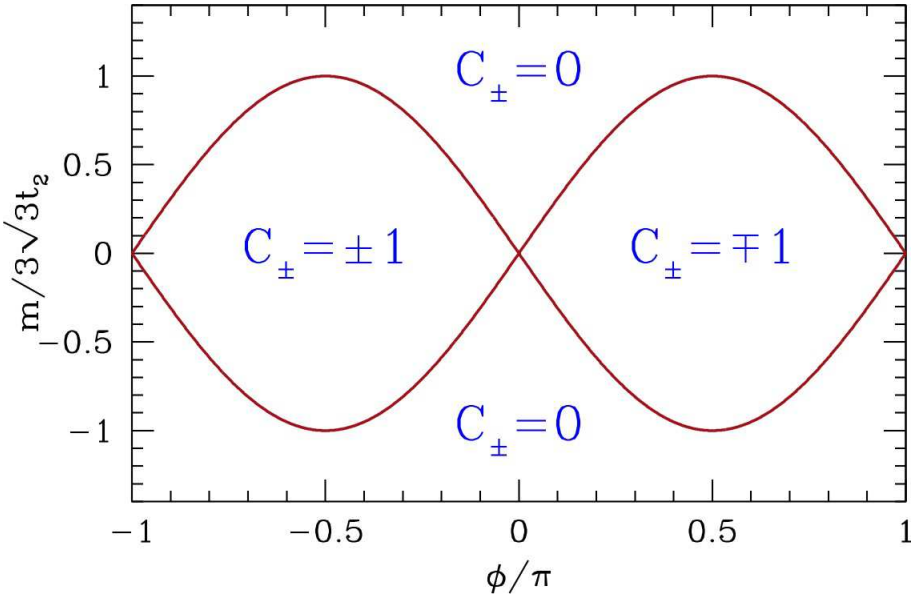


Figure 1.21: Topological phase diagram for the Haldane honeycomb lattice model, in which regions in the $(\sin\phi, m/t_2)$ cylinder are labeled by the Chern numbers $C_\pm$ of the energy bands.

There are three ways to introduce a gap into the model, *i.e.* to gap out the dispersion at the $\boldsymbol{K}$ and $\boldsymbol{K}'$ points at the two inequivalent Brillouin zone corners:

1. The first way is by introducing a Semenoff mass term, which is of the form $V_{\rm S} = \Delta_{\rm S}\,\sigma^z$. This turns graphene into boron nitride (BN), distinguishing the local π-orbital energies on the B and N sites. One can check that

$$\mathcal{P}\,\sigma^z\,\mathcal{P}^{-1} = -\sigma^z \quad , \qquad \mathcal{T}\,\sigma^z\,\mathcal{T}^{-1} = +\sigma^z \quad , \tag{1.205}$$

   and therefore the Semenoff mass breaks parity and preserves time-reversal.

2. The second way comes from the Haldane honeycomb lattice model at $m = 0$, where $V_{\rm H} = \Delta_{\rm H}\,\sigma^z\tau^z$.

$$\mathcal{P}\,\sigma^z\tau^z\,\mathcal{P}^{-1} = +\sigma^z\tau^z \quad , \qquad \mathcal{T}\,\sigma^z\tau^z\,\mathcal{T}^{-1} = -\sigma^z\tau^z \quad . \tag{1.206}$$

   This term, the Haldane mass, preserves parity but breaks time-reversal. It leads to a topological band structure in which the bands are classified by nonzero Chern numbers.

3. The third way involves introducing the physical electron spin, and arises from spin-orbit effects. It essentially is described by two copies of the Haldane model, in which the up and down spin electrons couple oppositely to magnetic flux. This was first discussed by Kane and Mele[36], and is described by the perturbation $V_{\rm KM} = \Delta_{\rm KM}\,\sigma^z\tau^z s^z$. where $\boldsymbol{s}$ is the

[36]C. L. Kane and E. J. Mele, *Phys. Rev. Lett.* **95**, 226801 (2005).

electron spin operator . The Kane-Mele mass term preserves $\mathcal{P}$ and $\mathcal{T}$ symmetries:

$$\mathcal{P}\,\sigma^z\tau^z s^z\,\mathcal{P}^{-1} = +\sigma^z\tau^z s^z \quad , \quad \mathcal{T}\,\sigma^z\tau^z s^z\,\mathcal{T}^{-1} = +\sigma^z\tau^z s^z \quad . \tag{1.207}$$

Therefore, following the tried and true rule in physics that "everything which is not forbidden is compulsory", there *must* be a KM mass term in real graphene. The catch is that it is extremely small because graphene is a low-Z atom, and first principles calculations[37] conclude that the spin-orbit gap is on the order of $10\,\mathrm{mK}$ – too small to be observed due to finite temperature and disorder effects. However, there are many materials (Bi bilayers, HgTe/CdTe heterostructures, various three-dimensional materials such as α-Sn, $\mathrm{Bi}_x\mathrm{Sb}_{1-x}$ and others) where the effect is predicted to be sizable and where it is indeed observed. This is the essence of topological insulator behavior.

1.7.4 The TKNN formula

Recall the Hamiltonian of Eqn. 1.163 for the isotropic square lattice Hofstadter model with flux $\phi = 2\pi p/q$ per unit cell. A more general version, incorporating anisotropy which breaks 90° rotational symmetry, is given by[38]

$$H(\theta_1,\theta_2) = -\begin{pmatrix} 2t_2\cos\theta_2 & t_1 & 0 & \cdots & 0 & t_1\,e^{-i\theta_1} \\ t_1 & 2t_2\cos\left(\theta_2+\frac{2\pi p}{q}\right) & t_1 & & & 0 \\ 0 & t_1 & 2t_2\cos\left(\theta_2+\frac{4\pi p}{q}\right) & t_1 & & \vdots \\ \vdots & 0 & t_1 & \ddots & & \vdots \\ & & & & & t_1 \\ 0 & & & \cdots & t_1 & 2t_2\cos\left(\theta_2+\frac{2\pi(q-1)p}{q}\right) \\ t_1\,e^{i\theta_1} & 0 & & & & \end{pmatrix} . \tag{1.208}$$

This is a $q\times q$ matrix, and the q eigenvectors $|\,\varphi_n(\boldsymbol{\theta})\,\rangle$ are labeled by a band index $n\in\{1,\ldots,q\}$, with component amplitudes $\varphi_{a,n}(\boldsymbol{\theta})$ satisfying

$$H_{aa'}(\boldsymbol{\theta})\,\varphi_{a',n}(\boldsymbol{\theta}) = E_n(\boldsymbol{\theta})\,\varphi_{a,n}(\boldsymbol{\theta}) \quad . \tag{1.209}$$

From Wigner-von Neumann, we expect generically that neighboring bands will not cross as a function of the two parameters (θ_1,θ_2), because degeneracy has codimension three. Thus, associated with each band n is a Chern number C_n. By color coding each spectral gap according to the Chern number of all bands below it, J. Avron produced a beautiful and illustrative image of Hofstadter's butterfly, shown in Fig. 1.22 for the isotropic square lattice and in Fig. 1.23 for the isotropic honeycomb lattice.

[37] See Y. Yao *et al.*, *Phys. Rev. B* **75**, 041401(R) (2007).

[38] We drop the hat on $\hat{H}(\boldsymbol{\theta})$ but fondly recall that $\hat{H}_{aa'}(\boldsymbol{\theta})$ is the lattice Fourier transform of $H_{aa'}(\boldsymbol{R}-\boldsymbol{R}')$.

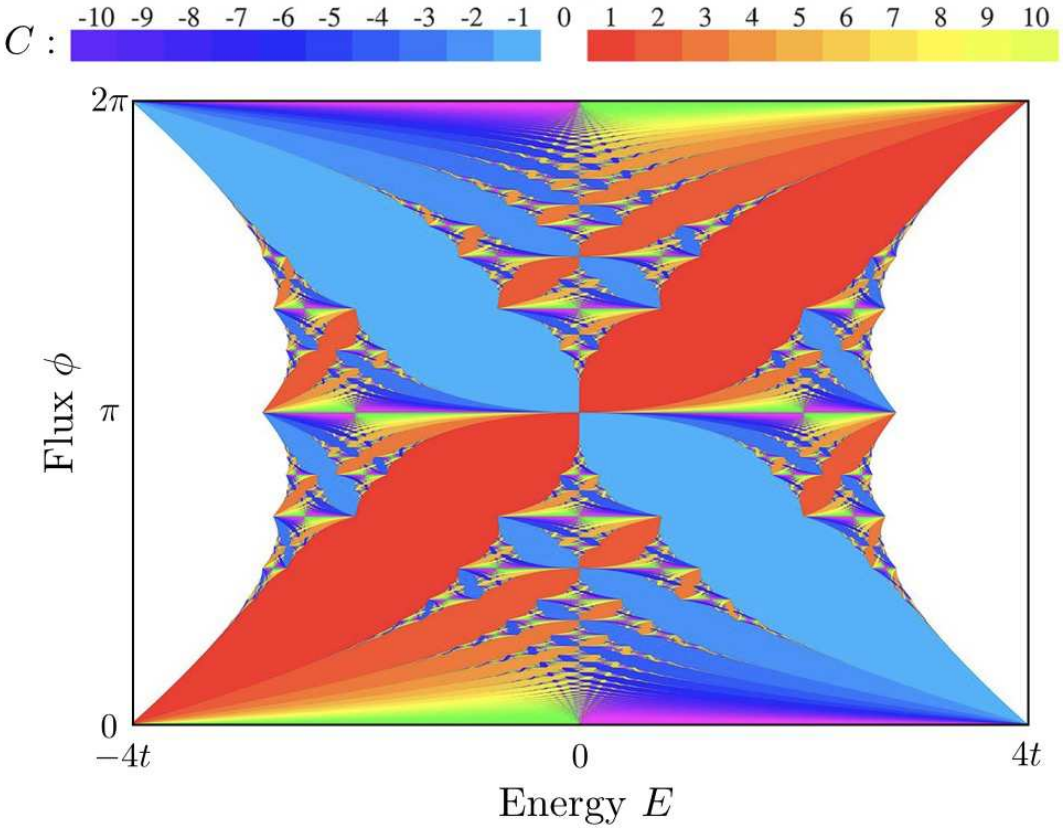


Figure 1.22: Avron's depiction of the Hofstadter butterfly for the isotropic square lattice system. The band gap regions are color coded by Chern number, C, which is the sum of the Chern numbers of all bands below a given gap. White regions correspond to $C = 0$. See J. E. Avron, *Colored Hofstadter butterflies*, in *Multiscale Methods in Quantum Mechanics*, P. Blanchard and G. Dell'Antonio, eds. (Birkhäuser, 2004).

It turns out that the Chern number is not just an abstract topological index. It is in fact the dimensionless Hall conductivity σ_{xy} itself, provided the Fermi level lies in a gap between magnetic subbands. This was first discovered by Thouless, Kohmoto, Nightingale, and den Nijs, in a seminal paper known by its authors' initials, TKNN[39]. In fact, we've developed the theory here in reverse chronological order. First came TKNN, who found that the contribution $\sigma_{xy}^{(n)}$ to the total Hall conductivity from a band lying entirely below the Fermi level is given by $\sigma_{xy}^{(n)} = \frac{e^2}{h}\, C_n$, where

$$C_n = \frac{i}{2\pi} \int_0^{2\pi} d\theta_1 \int_0^{2\pi} d\theta_2 \left(\left\langle \frac{\partial \varphi_n}{\partial \theta_1} \middle| \frac{\partial \varphi_n}{\partial \theta_2} \right\rangle - \left\langle \frac{\partial \varphi_n}{\partial \theta_2} \middle| \frac{\partial \varphi_n}{\partial \theta_1} \right\rangle \right) \tag{1.210}$$

is an integral over the Brillouin zone. They proved that this expression is an integer, because

[39] D. J. Thouless, M. Kohmoto, M. P. Nightingale, and M. den Nijs, *Phys. Rev. Lett.* **49**, 405 (1982).

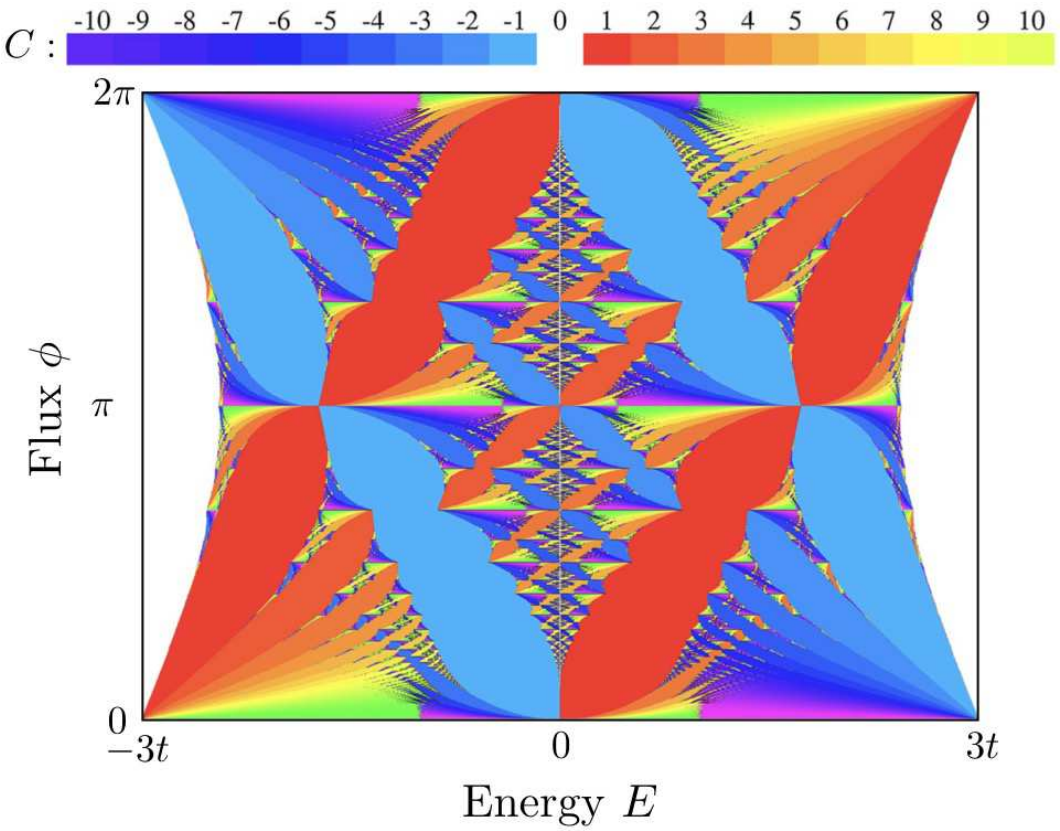


Figure 1.23: Colored Hofstadter butterfly for the honeycomb lattice system, from A. Agazzi, J.-P. Eckman, and G. M. Graf, *J. Stat. Phys.* **156**, 417 (2014).

invoking Stokes' theorem,

$$C_n = \frac{i}{2\pi}\int\limits_0^{2\pi} d\theta_2 \left\langle\, \varphi_n \,\middle|\, \frac{\partial\varphi_n}{\partial\theta_2} \,\right\rangle\Big|_{\theta_1=0}^{\theta_1=2\pi} - \frac{i}{2\pi}\int\limits_0^{2\pi} d\theta_1 \left\langle\, \varphi_n \,\middle|\, \frac{\partial\varphi_n}{\partial\theta_1} \,\right\rangle\Big|_{\theta_2=0}^{\theta_2=2\pi} \quad . \tag{1.211}$$

But since $H(\theta_1,\theta_2)$ is doubly periodic with period 2π in each of its arguments, we must have

$$\begin{aligned} |\,\varphi_n(\theta_1,2\pi)\,\rangle &= e^{if_n(\theta_1)}\,|\,\varphi_n(\theta_1,0)\,\rangle \\ |\,\varphi_n(2\pi,\theta_2)\,\rangle &= e^{ig_n(\theta_2)}\,|\,\varphi_n(0,\theta_2)\,\rangle \quad . \end{aligned} \tag{1.212}$$

Thus, one finds

$$C_n = \frac{1}{2\pi}\Big(f_n(2\pi) - f_n(0) + g_n(0) - g_n(2\pi)\Big) \quad . \tag{1.213}$$

But we also have

$$\begin{aligned} |\,\varphi_n(0,0)\,\rangle &= e^{-if_n(0)}\,|\,\varphi_n(0,2\pi)\,\rangle = e^{-if_n(0)}\,e^{-ig_n(2\pi)}\,|\,\varphi_n(2\pi,2\pi)\,\rangle \\ &= e^{-if_n(0)}\,e^{-ig_n(2\pi)}\,e^{if_n(2\pi)}\,|\,\varphi_n(2\pi,0)\,\rangle = e^{-if_n(0)}\,e^{-ig_n(2\pi)}\,e^{if_n(2\pi)}\,e^{ig_n(0)}\,|\,\varphi_n(0,0)\,\rangle \quad , \end{aligned} \tag{1.214}$$

and therefore $\exp(2\pi i\, C_n) = 1$ and $C_n \in \mathbb{Z}$. But just as Berry didn't know he had found a holonomy, TKNN didn't know they had found a Chern number. That mathematical feature was first elucidated by Avron, Seiler, and Simon[40], in a paper which is widely appreciated but which, understandably, is not known by its authors' initials.

To see why Hall conductivity is related to Berry curvature, consider an electric field $\boldsymbol{E} = E_y\,\hat{\boldsymbol{y}}$, and the single electron Hamiltonian $H(E_y) = H(0) - eE_y y$, where $H(0) = \frac{\boldsymbol{\pi}^2}{2m} + V(\boldsymbol{r})$ has eigenstates $|\,\alpha\,\rangle$ and eigenvalues ε_α. First order perturbation theory in the electric field term says

$$|\,\alpha'\,\rangle = |\,\alpha\,\rangle - eE_y \sum_\beta{}' \frac{|\,\beta\,\rangle\langle\,\beta\,|\,y\,|\,\alpha\,\rangle}{\varepsilon_\alpha - \varepsilon_\beta} \quad , \tag{1.215}$$

where the prime on the sum means the term with $\beta = \alpha$ is excluded. Let's now compute the expectation of the velocity operator v_x in the perturbed state $|\,\alpha'\,\rangle$. We have, to lowest order,

$$\langle\,\alpha'\,|\,v_x\,|\,\alpha'\,\rangle = -eE_y \sum_\beta{}' \frac{\langle\,\alpha\,|\,v_x\,|\,\beta\,\rangle\langle\,\beta\,|\,y\,|\,\alpha\,\rangle + \langle\,\alpha\,|\,y\,|\,\beta\,\rangle\langle\,\beta\,|\,v_x\,|\,\alpha\,\rangle}{\varepsilon_\alpha - \varepsilon_\beta} \tag{1.216}$$

We now invoke the Feynman-Hellman theorem, which says

$$\langle\,\alpha\,|\,y\,|\,\beta\,\rangle = \frac{\hbar}{i}\,\frac{\langle\,\alpha\,|\,v_y\,|\,\beta\,\rangle}{\varepsilon_\alpha - \varepsilon_\beta} \quad , \tag{1.217}$$

multiply by the electron charge $-e$, divide by the area of the system Ω, and sum using the Fermi distribution over the levels $|\,\alpha\,\rangle$, to obtain the current density j_x :

$$j_x = E_y \cdot \frac{e^2}{h} \cdot \frac{2\pi i\hbar^2}{\Omega} \sum_\alpha \sum_\beta f_\alpha\,(1 - f_\beta)\,\epsilon_{ij}\,\frac{\langle\,\alpha\,|\,v_i\,|\,\beta\,\rangle\langle\,\beta\,|\,v_j\,|\,\alpha\,\rangle}{(\varepsilon_\alpha - \varepsilon_\beta)^2} \quad , \tag{1.218}$$

where f_α is the Fermi function at temperature T, chemical potential μ, and energy ε_α. The above expression for $\sigma_{xy} = j_x/E_y$ is known as the *Kubo formula* for the Hall conductivity. At $T = 0$, the Fermi distribution becomes the step function $f_\alpha = \Theta(E_{\mathsf{F}} - E_\alpha)$.

Suppose our system lies on a torus defined by the spatial periods $\boldsymbol{L}_1$ and $\boldsymbol{L}_2$. Define the gauge transformed Hamiltonian

$$\widetilde{H}(\boldsymbol{\theta}) \equiv e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\,H\,e^{+i\boldsymbol{q}\cdot\boldsymbol{r}} \quad , \tag{1.219}$$

where

$$\boldsymbol{q} = \theta_2\,\frac{\hat{\boldsymbol{z}}\times\boldsymbol{L}_1}{\Omega} - \theta_1\,\frac{\hat{\boldsymbol{z}}\times\boldsymbol{L}_2}{\Omega} \quad , \tag{1.220}$$

with $\Omega = \hat{\boldsymbol{z}}\cdot\boldsymbol{L}_1\times\boldsymbol{L}_2 = 2\pi\ell^2 p$ with $p \in \mathbb{Z}$, *i.e.* the total magnetic flux through the system is an integer multiple of the Dirac quantum. Then

$$\frac{\partial\widetilde{H}}{\partial\theta_i} = \frac{\partial\boldsymbol{q}}{\partial\theta_i}\cdot e^{-i\boldsymbol{q}\cdot\boldsymbol{r}}\,\hbar\boldsymbol{v}\,e^{i\boldsymbol{q}\cdot\boldsymbol{r}} \equiv \frac{\partial\boldsymbol{q}}{\partial\theta_i}\cdot\hbar\widetilde{\boldsymbol{v}} \quad , \tag{1.221}$$

[40] J. E. Avron, R. Seiler, and B. Simon, *Phys. Rev. Lett.* **51**, 51 (1983).

because $[H, \boldsymbol{r}] = (\hbar/i)\,\boldsymbol{v}$. Thus, defining $|\,\widetilde{\alpha}\,\rangle \equiv \exp(-i\boldsymbol{q}\cdot\boldsymbol{r})\,|\,\alpha\,\rangle$, and recalling the definition of the wavevector $\boldsymbol{q} = \epsilon_{ab}\,\theta_a\,\boldsymbol{L}_b \times \hat{\boldsymbol{z}}/\Omega$, we find

$$\frac{\partial\widetilde{H}}{\partial\theta_a} = \hbar\,\epsilon_{ab}\,\epsilon_{ij}\,\frac{\widetilde{v}_i\,L_b^j}{\Omega} \quad . \tag{1.222}$$

We then find

$$\sigma_{xy} = \frac{j_x}{E_y} = \sum_{\alpha\ \mathrm{occ}} \sigma_{xy}^{(\alpha)} \quad , \tag{1.223}$$

where the sum is over occupied states below the Fermi level, and where

$$\sigma_{xy}^{(\alpha)} = \frac{e^2}{h}\cdot 2\pi i \sum_\beta{}' \epsilon_{ij}\,\frac{\langle\,\widetilde{\alpha}\,|\,\frac{\partial\widetilde{H}}{\partial\theta_i}\,|\,\widetilde{\beta}\,\rangle\langle\,\widetilde{\beta}\,|\,\frac{\partial\widetilde{H}}{\partial\theta_j}\,|\,\widetilde{\alpha}\,\rangle}{(\varepsilon_\alpha - \varepsilon_\beta)^2} \quad , \tag{1.224}$$

which is precisely of the form of Eqn. 1.183. Thus, if we now *uniformly average over the boundary phases* θ_1 *and* θ_2, we obtain

$$\big\langle \sigma_{xy}^{(\alpha)} \big\rangle = \frac{e^2}{h}\cdot\frac{i}{2\pi}\int_0^{2\pi} d\theta_1 \int_0^{2\pi} d\theta_2 \sum_\beta{}' \epsilon_{ij}\,\frac{\langle\,\widetilde{\alpha}\,|\,\frac{\partial\widetilde{H}}{\partial\theta_i}\,|\,\widetilde{\beta}\,\rangle\langle\,\widetilde{\beta}\,|\,\frac{\partial\widetilde{H}}{\partial\theta_j}\,|\,\widetilde{\alpha}\,\rangle}{(\varepsilon_\alpha - \varepsilon_\beta)^2} = \frac{e^2}{h}\,C_\alpha \quad , \tag{1.225}$$

i.e. each filled band α contributes $\frac{e^2}{h}\,C_\alpha$ to the total Hall conductivity whenever the Fermi level at $T = 0$ lies in a gap between energy bands. For a crystalline (periodic) system, averaging over $\theta_{1,2}$ is tantamount to integrating over the Brillouin zone.

1.8 Appendix I : Basis Wavefunctions on a Torus

Periodic boundary conditions on the torus requires $t(\boldsymbol{L}_a)\,|\,\psi\,\rangle = e^{i\theta_a}\,|\,\psi\,\rangle$ for all states $|\,\psi\,\rangle$. Let's examine what this requires for the analytic part $f(z)$. We have

$$\begin{aligned} t(\boldsymbol{L}) &= e^{i\boldsymbol{\kappa}\cdot\boldsymbol{L}/\hbar} = e^{i(\kappa\bar{L}+\kappa^\dagger L)/2\hbar} \\ &= e^{-L\bar{L}/4\ell^2}\,e^{i\kappa\bar{L}/2\hbar}\,e^{i\kappa^\dagger L/2\hbar} \\ &= e^{-L\bar{L}/4\ell^2}\,e^{i\chi}\,e^{z\bar{z}/4\ell^2}\,e^{\bar{L}\bar{\partial}}\,e^{-z\bar{z}/2\ell^2}\,e^{L\partial}\,e^{z\bar{z}/4\ell^2}\,e^{-i\chi} \quad . \end{aligned} \tag{1.226}$$

Thus, with $\psi(\boldsymbol{r}) = e^{i\chi}\,e^{-z\bar{z}/4\ell^2}\,f(z)$, we have

$$\begin{aligned} t(\boldsymbol{L})\,\psi(\boldsymbol{r}) &= e^{-L\bar{L}/4\ell^2}\,e^{i\chi}\,e^{z\bar{z}/4\ell^2}\,e^{\bar{L}\bar{\partial}}\,e^{-z\bar{z}/2\ell^2}\,e^{L\partial}\,f(z) \\ &= e^{i\chi}\,e^{-L\bar{L}/4\ell^2}\,e^{-z\bar{z}/4\ell^2}\,e^{-z\bar{L}/2\ell^2}\,f(z+L) \quad . \end{aligned} \tag{1.227}$$

Thus, we must have

$$f(z+L_a) = e^{i\theta_a}\,e^{L_a\bar{L}_a/4\ell^2}\,e^{z\bar{L}_a/2\ell^2}\,f(z) \quad , \tag{1.228}$$

valid for $a = 1, 2$. Note that integrating the logarithmic derivative of $f(z)$ around the torus yields

$$\begin{aligned}\oint\limits_{\Omega} \frac{dz}{2\pi i}\ \frac{f'(z)}{f(z)} &= \int\limits_0^{L_1} \frac{dz}{2\pi i}\ \frac{d}{dz}\ \ln\Big(\frac{f(z)}{f(z+L_2)}\Big) + \int\limits_0^{L_2} \frac{dz}{2\pi i}\ \frac{d}{dz}\ \ln\Big(\frac{f(z+L_1)}{f(z)}\Big) \\ &= -\int\limits_0^{L_1} \frac{dz}{2\pi i}\ \frac{\bar{L}_2}{2\ell^2} + \int\limits_0^{L_2} \frac{dz}{2\pi i}\ \frac{\bar{L}_1}{2\ell^2} = \frac{\bar{L}_1 L_2 - L_1 \bar{L}_2}{2\pi i} \cdot \frac{1}{2\ell^2} = \frac{\Omega}{2\pi\ell^2} = N_\phi \quad ,\end{aligned} \tag{1.229}$$

which establishes that $f(z)$ has precisely N_ϕ zeros on the torus.

Now consider the Jacobi theta function,

$$\vartheta_1(w\,|\,\sigma) = -i \sum_{n=-\infty}^{\infty} (-1)^n\, e^{i\pi\sigma(n+\frac{1}{2})^2}\, e^{(2n+1)iw} \quad , \tag{1.230}$$

where $\mathsf{Im}\,(\sigma) > 0$. This function is quasiperiodic over the fundamental domain for w, which is a parallelogram with sides 1 and σ, satisfying

$$\begin{aligned}\vartheta_1(w+\pi\,|\,\sigma) &= -\vartheta_1(w\,|\,\sigma) \\ \vartheta_1(w+\pi\sigma\,|\,\sigma) &= -e^{-2iw}\, e^{-i\pi\sigma}\, \vartheta_1(w\,|\,\sigma) \quad .\end{aligned} \tag{1.231}$$

From the above relations, integrating the logarithmic derivative of $\vartheta_1(w\,|\,\sigma)$ establishes that the function has one zero in the fundamental domain, located at $w = 0$. Iterating the second of the above relations, one has

$$\vartheta_1(w + j\pi\sigma\,|\,\sigma) = (-1)^j\, e^{-2ijw}\, e^{-ij^2\pi\sigma}\, \vartheta_1(w\,|\,\sigma) \quad , \tag{1.232}$$

for any integer j. In addition, using the Poisson summation formula,

$$\sum_{n=-\infty}^{\infty} \delta(x-n) = \sum_{m=-\infty}^{\infty} e^{2\pi i m x} \quad , \tag{1.233}$$

one derives the identity

$$\vartheta_1(w\,|\,\sigma) = \sqrt{\tfrac{i}{\sigma}}\, e^{-i\pi(1+\sigma)/4}\, e^{-iw^2/\pi\sigma}\, \vartheta\big(w\,\big|\,-\tfrac{1}{\sigma}\big) \quad . \tag{1.234}$$

Now consider the function

$$f(z) = e^{\bar{L}_1 z^2/4\ell^2 L_1}\, e^{i\lambda z/L_1}\, \vartheta_1\bigg(\frac{\pi z}{L_1} - \pi\zeta\,\bigg|\,\frac{L_2}{N_\phi\, L_1}\bigg) \quad . \tag{1.235}$$

Then

$$\begin{aligned} f(z+L_1) &= -e^{L_1\bar{L}_1/4\ell^2}\,e^{\bar{L}_1 z/2\ell^2}\,e^{i\lambda}\,f(z) \\ f(z+L_2) &= (-1)^{N_\phi}\,e^{L_2\bar{L}_2/4\ell^2}\,e^{\bar{L}_2 z/2\ell^2}\,e^{i\lambda\tau}\,e^{2\pi i N_\phi\zeta}\,f(z) \quad , \end{aligned} \tag{1.236}$$

where $\tau \equiv L_2/L_1$. Invoking the periodicity requirements, we obtain

$$\begin{aligned} e^{i\lambda} &= -e^{i\theta_1} \\ e^{i\lambda\tau}\,e^{2\pi i N_\phi\zeta} &= (-1)^{N_\phi}\,e^{i\theta_2} \quad . \end{aligned} \tag{1.237}$$

Thus, we have

$$\lambda = \theta_1 + (2k_1+1)\pi \tag{1.238}$$

and

$$\zeta = \frac{\theta_2 + \pi N_\phi + 2\pi k_2}{2\pi N_\phi} - \frac{\theta_1 + (2k_1+1)\pi}{2\pi N_\phi}\cdot\tau \quad , \tag{1.239}$$

where k_1 and k_2 are integers. Since $k_1 \to k_1 + 1$ increases the argument of the ϑ-function by a multiple of $\sigma \equiv \tau/N_\phi$, one can invoke the quasiperiodicity relation, whence one finds that this results in a multiplication of $f(z)$ by a constant. Therefore, we are free to select a fixed value for k_1. We choose $k_1 \equiv -1$. Then

$$\begin{aligned} \lambda &= \theta_1 - \pi \\ \zeta &= \frac{\theta_2 + \pi N_\phi + 2\pi k_2}{2\pi N_\phi} + \frac{(\pi-\theta_1)\,\tau}{2\pi N_\phi} \quad . \end{aligned} \tag{1.240}$$

So our basis functions are

$$\psi_k(\boldsymbol{r}) = C\,e^{i\chi}\,e^{i\pi k/N_\phi}\,e^{-z\bar{z}/4\ell^2}\,e^{\bar{L}_1 z^2/4\ell^2 L_1}\,e^{i(\theta_1-\pi)z/L_1}\,\vartheta_1\!\left(\frac{\pi z}{L_1} - \pi\zeta_k \,\middle|\, \frac{L_2}{N_\phi\,L_1}\right) \quad , \tag{1.241}$$

where

$$\zeta_k = \frac{\theta_2 + \pi N_\phi + 2\pi k}{2\pi N_\phi} + \frac{(\pi-\theta_1)\,\tau}{2\pi N_\phi} \tag{1.242}$$

and C is a constant independent of k. Then after some work one can show that indeed

$$\begin{aligned} t_1\,\psi_k(\boldsymbol{r}) &= e^{i\theta_1/N_\phi}\,\psi_{k-1}(\boldsymbol{r}) \\ t_2\,\psi_k(\boldsymbol{r}) &= e^{i\theta_2/N_\phi}\,e^{2\pi i k/N_\phi}\,\psi_k(\boldsymbol{r}) \quad . \end{aligned} \tag{1.243}$$

Finally, define

$$w \equiv \frac{z}{L_1} \equiv u + \tau v \quad , \tag{1.244}$$

where $(u,v) \in [0,1] \times [0,1]$. Then with $\phi_x(\boldsymbol{w}) \equiv \psi_k(z = wL_1, \bar{z} = \bar{w}\bar{L}_1)$,

$$\begin{aligned}\phi_k(\boldsymbol{w}) &= C\, e^{i\pi k/N_\phi}\, e^{\pi N_\phi (w-\bar{w})/2\tau_2}\, e^{i(\theta_1-\pi)w}\, \vartheta_1\Big(\pi w - \pi\zeta_k \,\Big|\, \frac{\tau}{N_\phi}\Big) \\ &= C\, e^{i\pi k/N_\phi}\, e^{i\pi N_\phi(uv+\tau v^2)}\, e^{i(\theta_1-\pi)(u+\tau v)}\, \vartheta_1\Big(\pi u - \frac{k\pi}{N_\phi} - \frac{\theta_2+\pi N_\phi}{2N_\phi} + \pi v\tau + \frac{(\theta_1-\pi)\,\tau}{2N_\phi} \,\Big|\, \frac{\tau}{N_\phi}\Big) \quad ,\end{aligned} \tag{1.245}$$

which is holomorphic in τ. The normalization condition is

$$1 = \Omega \int_0^1 du \int_0^1 dv\, \big|\phi_k(u,v)\big|^2 = |C|^2 \left(\frac{2\pi\ell^2 N_\phi^3}{\tau_2}\right)^{1/2} \exp\left(\frac{(\theta_1-\pi)^2\,\tau_2}{2\pi N_\phi}\right) \quad . \tag{1.246}$$

1.9 Appendix II : Coherent States and their Path Integral

1.9.1 Feynman path integral

The path integral formulation of quantum mechanics is both beautiful and powerful. It is useful in elucidating the quantum-classical correspondence and the semiclassical approximation, in accounting for interference effects, in treatments of tunneling problems via the method of instantons, *etc.* Our goal is to derive and to apply a path integral method for quantum spin. We begin by briefly reviewing the derivation of the usual Feynman path integral.

Consider the propagator $K(x_{\rm i}, x_{\rm f}, T)$, which is the probability amplitude that a particle located at $x = x_{\rm i}$ at time $t = 0$ will be located at $x = x_{\rm f}$ at time $t = T$. We may write

$$\begin{aligned}K(x_{\rm i}, x_{\rm f}, T) &= \langle\, x_{\rm f} \,|\, e^{-iHT/\hbar} \,|\, x_{\rm i} \,\rangle \\ &= \langle\, x_N \,|\, e^{-i\epsilon H/\hbar}\, \mathbf{1}\, e^{-i\epsilon H/\hbar}\, \mathbf{1} \cdots \mathbf{1}\, e^{-i\epsilon H/\hbar} \,|\, x_0 \,\rangle\end{aligned} \tag{1.247}$$

where $\epsilon = T/N$, and where we have defined $x_0 \equiv x_{\rm i}$ and $x_N \equiv x_{\rm f}$. We are interested in the limit $N \to \infty$. Inserting $(N-1)$ resolutions of the identity of the form

$$\mathbf{1} = \int_{-\infty}^{\infty} dx_j \,|\, x_j \,\rangle\langle\, x_j \,| \quad , \tag{1.248}$$

we find that we must evaluate matrix elements of the form

$$\begin{aligned}\langle\, x_{j+1} \,|\, e^{-iH\epsilon/\hbar} \,|\, x_j \,\rangle &\approx \int_{-\infty}^{\infty} dp_j\, \langle\, x_{j+1} \,|\, p_j \,\rangle\, \langle\, p_j \,|\, e^{-iT\epsilon/\hbar}\, e^{-iV\epsilon/\hbar} \,|\, x_j \,\rangle \\ &= \int_{-\infty}^{\infty} dp_j\, e^{ip_j(x_{j+1}-x_j)}\, e^{-i\epsilon p_j^2/2m\hbar}\, e^{-i\epsilon V(x_j)/\hbar} \quad .\end{aligned} \tag{1.249}$$

The propagator may now be written as

$$\begin{aligned}\langle\, x_N \,|\, e^{-iHT/\hbar} \,|\, x_0 \,\rangle &\approx \int\limits_{-\infty}^{\infty} \prod_{j=1}^{N-1} dx_j \int\limits_{-\infty}^{\infty} \prod_{k=0}^{N-1} dp_k \, \exp\left\{ i \sum_{k=0}^{N-1} \left[p_k(x_{k+1} - x_k) - \frac{\epsilon\, p_k^2}{2m\hbar} - \frac{\epsilon}{\hbar}\, V(x_k) \right] \right\} \\ &= \left(\frac{2\pi\hbar m}{i\epsilon} \right)^N \int\limits_{-\infty}^{\infty} \prod_{j=1}^{N-1} dx_j \, \exp\left\{ \frac{i\epsilon}{\hbar} \sum_{k=1}^{N-1} \left[\tfrac{1}{2} m \Big(\frac{x_{j+1} - x_j}{\epsilon} \Big)^2 - V(x_j) \right] \right\} \\ &\equiv \int\limits_{\substack{x(0)=x_{\rm i} \\ x(T)=x_{\rm f}}} \mathcal{D}x(t) \, \exp\left\{ \frac{i}{\hbar} \int_0^T dt \left[\tfrac{1}{2} m\dot{x}^2 - V(x) \right] \right\} \quad , \end{aligned} \tag{1.250}$$

where we absorb the prefactor into the measure $\mathcal{D}x(t)$. Note the boundary conditions on the path integral at $t = 0$ and $t = T$. In the semiclassical approximation, we assume that the path integral is dominated by trajectories $x(t)$ which extremize the argument of the exponential in the last term above. This quantity is (somewhat incorrectly) identified as the classical action, $\mathcal{S}$, and the action-extremizing equations are of course the Euler-Lagrange equations. Setting $\delta\mathcal{S} = 0$ yields Newton's second law, $m\ddot{x} = -\partial V/\partial x$, which is to be solved subject to the two boundary conditions.

The imaginary time version, which yields the thermal propagator, is obtained by writing $T = -i\hbar\beta$ and $t = -i\tau$, in which case

$$\langle\, x_{\rm f} \,|\, e^{-\beta H} \,|\, x_{\rm i} \,\rangle = \int\limits_{\substack{x(0)=x_{\rm i} \\ x(\hbar\beta)=x_{\rm f}}} \mathcal{D}x(\tau) \, \exp\left\{ -\frac{1}{\hbar} \overbrace{\int_0^{\hbar\beta} d\tau \left[\tfrac{1}{2} m\dot{x}^2 + V(x) \right]}^{\text{Euclidean action } \mathcal{S}_{\rm E}} \right\} \quad . \tag{1.251}$$

The partition function is the trace of the thermal propagator, *viz.*

$$Z = \mathsf{Tr}\; e^{-\beta H} = \int\limits_{-\infty}^{\infty} dx \, \langle\, x \,|\, e^{-\beta H} \,|\, x \,\rangle = \int\limits_{x(0)=x(\hbar\beta)} \mathcal{D}x(\tau) \, \exp\big(-\mathcal{S}_{\rm E}[x(\tau)]/\hbar \big) \tag{1.252}$$

The equations of motion derived from $\mathcal{S}_{\rm E}$ are $m\ddot{x} = +\partial V/\partial x$, corresponding to motion in the 'inverted potential'.

1.9.2 Primer on coherent states

We now turn to the method of coherent state path integration. In order to discuss this, we must first introduce the notion of coherent states. This is most simply done by appealing to the

one-dimensional simple harmonic oscillator,

$$H = \frac{p^2}{2m} + \tfrac{1}{2}m\omega_0^2 x^2 = \hbar\omega_0\,(a^\dagger a + \tfrac{1}{2}) \tag{1.253}$$

where a and $a^\dagger$ are ladder operators,

$$a = \ell\,\partial_x + \frac{x}{2\ell} \qquad , \qquad a^\dagger = -\ell\,\partial_x + \frac{x}{2\ell} \tag{1.254}$$

with $\ell \equiv \sqrt{\hbar/2m\omega_0}$. *Exercise: Check that* $[a, a^\dagger] = 1$.

The ground state satisfies $a\,\psi_0(x) = 0$, which yields

$$\psi_0(x) = (2\pi\ell^2)^{-1/4}\,\exp(-x^2/4\ell^2) \quad . \tag{1.255}$$

The normalized coherent state $|\,z\,\rangle$ is defined as

$$|\,z\,\rangle = e^{-\frac{1}{2}|z|^2}\,e^{za^\dagger}|\,0\,\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,|\,n\,\rangle \quad . \tag{1.256}$$

The coherent state is an eigenstate of the annihilation operator a:

$$a\,|\,z\,\rangle = z\,|\,z\,\rangle \qquad \Longleftrightarrow \qquad \langle\,z\,|\,a^\dagger = \langle\,z\,|\,\bar{z} \quad . \tag{1.257}$$

The overlap of coherent states is given by

$$\langle\,z_1\,|\,z_2\,\rangle = e^{-\frac{1}{2}|z_1|^2}\,e^{-\frac{1}{2}|z_2|^2}\,e^{\bar{z}_1 z_2} \quad , \tag{1.258}$$

hence different coherent states are not orthogonal. Despite this nonorthogonality, the coherent states allow a simple resolution of the identity,

$$\mathbf{1} = \int\frac{d^2z}{2\pi i}\,|\,z\,\rangle\langle\,z\,| \qquad , \qquad \frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}\,z\;d\,\mathrm{Im}\,z}{\pi} \tag{1.259}$$

which is straightforward to establish.

To gain some physical intuition about the coherent states, define

$$z \equiv \frac{Q}{2\ell} + \frac{i\ell P}{\hbar} \quad . \tag{1.260}$$

One finds (*exercise!*)

$$\psi_{P,Q}(x) = \langle\,x\,|\,z\,\rangle = (2\pi\ell^2)^{-1/4}\,e^{-iPQ/2\hbar}\,e^{iPx/\hbar}\,e^{-(x-Q)^2/4\ell^2} \quad , \tag{1.261}$$

hence the coherent state $\psi_{P,Q}(x)$ is a wavepacket Gaussianly localized about $x = Q$, but oscillating with momentum P.

For example, we can compute

$$\begin{aligned}\langle\, Q,P\,|\,q\,|\,Q,P\,\rangle &= \langle\, z\,|\,\ell\,(a+a^\dagger)\,|\,z\,\rangle = 2\ell\ \mathsf{Re}\ z = Q \\ \langle\, Q,P\,|\,p\,|\,Q,P\,\rangle &= \langle\, z\,|\,\frac{\hbar}{2i\ell}\,(a-a^\dagger)\,|\,z\,\rangle = \frac{\hbar}{\ell}\,\mathsf{Im}\ z = P\end{aligned} \tag{1.262}$$

as well as

$$\begin{aligned}\langle\, Q,P\,|\,q^2\,|\,Q,P\,\rangle &= \langle\, z\,|\,\ell^2\,(a+a^\dagger)^2\,|\,z\,\rangle = Q^2+\ell^2 \\ \langle\, Q,P\,|\,p^2\,|\,Q,P\,\rangle &= -\langle\, z\,|\,\frac{\hbar^2}{4\ell^2}(a-a^\dagger)^2\,|\,z\,\rangle = P^2+\frac{\hbar^2}{4\ell^2}\quad .\end{aligned} \tag{1.263}$$

Thus, the root mean square fluctuations in the coherent state $|\,Q,P\,\rangle$ are

$$\Delta q = \ell = \sqrt{\frac{\hbar}{2m\omega_0}}\qquad,\qquad \Delta p = \frac{\hbar}{2\ell} = \sqrt{\frac{m\hbar\omega_0}{2}}\quad, \tag{1.264}$$

and $\Delta q\cdot\Delta p = \frac{1}{2}\,\hbar$. Thus we learn that the coherent state $\psi_{Q,P}(q)$ is localized in phase space, *i.e.* in both position and momentum. If we have a general operator $\hat{A}(q,p)$, we can then write

$$\langle\, Q,P\,|\,\hat{A}(q,p)\,|\,Q,P\,\rangle = A(Q,P) + \mathcal{O}(\hbar)\quad, \tag{1.265}$$

where $A(Q,P)$ is formed from $\hat{A}(q,p)$ by replacing $q\to Q$ and $p\to P$. Since

$$\frac{d^2z}{2\pi i} \equiv \frac{d\,\mathsf{Re}\,z\,d\,\mathsf{Im}\,z}{\pi} = \frac{dQ\,dP}{2\pi\hbar}\quad, \tag{1.266}$$

we can write the trace using coherent states as

$$\mathsf{Tr}\,\hat{A} = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty}\!dQ\int\limits_{-\infty}^{\infty}\!dP\,\langle\, Q,P\,|\,\hat{A}\,|\,Q,P\,\rangle\quad. \tag{1.267}$$

We now can understand the origin of the factor $2\pi\hbar$ in the denominator of each (q_i,p_i) integral over classical phase space in the metric $d\mu = \prod_i \frac{dq_i dp_i}{2\pi\hbar}$.

Note that ω_0 is arbitrary in our discussion. By increasing ω_0, the states become more localized in q and more plane wave like in p. However, so long as ω_0 is finite, the width of the coherent state in each direction is proportional to $\hbar^{1/2}$, and thus vanishes in the classical limit.

The resolution of the identity in Eqn. 1.259 prompts the following question. Suppose we consider an infinite discrete lattice $\{|\,z_{m,n}\,\rangle\}$ of coherent states, with $z_{m,n} = (m+in)\sqrt{\pi}$, called a *von Neumann lattice* of coherent states. The dimensionless phase space area A per unit cell of this lattice is π, which is the denominator in the integral resolution of the identity in Eqn. 1.259. One might expect, then, that while this basis is not orthogonal, since

$$\langle\, z_{m,n}\,|\,z_{m',n'}\,\rangle = e^{-\pi(m-m')^2/2}\,e^{-\pi(n-n')^2/2}\,e^{i\pi(mn'-nm')}\quad, \tag{1.268}$$

that the Gaussian overlaps could be undone, *i.e.* the overlap matrix could be inverted, and a complete and orthonormal set of *localized* LLL wavefunctions could be constructed. Alas, this is impossible! As shown by Perelomov[41], there is a single linear dependence relation among the von Neumann lattice of coherent states, when $A = \pi$. For $S < \pi$, the lattice is overcomplete by a finite amount per unit area. Similarly, when $A > \pi$, the lattice is undercomplete by a finite amount per unit area. But when $A = \pi$, it is undercomplete by precisely *one state*. This state is necessary to include in order for the filled Landau level to carry a Chern number $C = 1$. Thus, the von Neumann lattice cannot be used as a basis to describe the quantum Hall effect.

1.9.3 Coherent state path integral

Now we derive the imaginary time path integral. We write

$$\langle\, z_{\rm f} \,|\, e^{-\beta H} \,|\, z_{\rm i} \,\rangle = \langle\, z_N \,|\, e^{-\epsilon H/\hbar}\, \mathbf{1}\, e^{-\epsilon H/\hbar} \cdots \mathbf{1}\, e^{-\epsilon H/\hbar} \,|\, z_0 \,\rangle \quad , \tag{1.269}$$

inserting resolutions of the identity at $N-1$ points, as before. We next evaluate the matrix element

$$\begin{aligned} \langle\, z_j \,|\, e^{-\epsilon H/\hbar} \,|\, z_{j-1} \,\rangle &= \langle\, z_j \,|\, z_{j-1} \,\rangle \cdot \left\{ 1 - \frac{\epsilon}{\hbar} \frac{\langle\, z_j \,|\, H \,|\, z_{j-1} \,\rangle}{\langle\, z_j \,|\, z_{j-1} \,\rangle} + \dots \right\} \\ &\simeq \langle\, z_j \,|\, z_{j-1} \,\rangle \, \exp\left\{ -\frac{\epsilon}{\hbar} H(\bar z_j \,|z_{j-1}) \right\} \end{aligned} \tag{1.270}$$

where

$$H(\bar z \,|w) \equiv \frac{\langle\, z \,|\, H \,|\, w \,\rangle}{\langle\, z \,|\, w \,\rangle} = e^{-\bar z w} \,\langle\, 0 \,|\, e^{\bar z a}\, H(a^\dagger, a)\, e^{w a^\dagger} \,|\, 0 \,\rangle \quad . \tag{1.271}$$

This last equation is extremely handy. It says, upon invoking eqn. 1.257, that if $H(a, a^\dagger)$ is *normal ordered* such that all creation operators $a^\dagger$ appear to the *left* of all destruction operators a, then $H(\bar z|w)$ is obtained from $H(a^\dagger, a)$ simply by sending $a^\dagger \to \bar z$ and $a \to w$. This is because a acting to the right on $|\, w \,\rangle$ yields its eigenvalue w, while $a^\dagger$ acting to the left on $\langle\, z \,|$ generates $\bar z$. Note that the function $H(\bar z|w)$ is holomorphic in w and in $\bar z$, but is completely independent of their complex conjugates $\bar w$ and z.

The overlap between coherent states at consecutive time slices may be written

$$\langle\, z_j \,|\, z_{j-1} \,\rangle = \exp\left\{ -\tfrac{1}{2}\Big[\bar z_j (z_j - z_{j-1}) - z_{j-1}(\bar z_j - \bar z_{j-1}) \Big] \right\} \quad , \tag{1.272}$$

[41] A. M. Perelomov, *Theor. Math. Phys.* **6**, 156 (1971).

hence

$$\begin{aligned}\langle\, z_N \,|\, z_{N-1} \,\rangle \cdots \langle\, z_1 \,|\, z_0 \,\rangle = \exp\left\{ \tfrac{1}{2} \sum_{j=1}^{N-1} \Big[z_j(\bar z_{j+1} - \bar z_j) - \bar z_j(z_j - z_{j-1}) \Big] \right\} \\ \times \exp\left\{ \tfrac{1}{2}\, z_0(\bar z_1 - \bar z_0) - \tfrac{1}{2}\, \bar z_N(z_N - z_{N-1}) \right\}\end{aligned} \tag{1.273}$$

which allows us to write down the path integral expression for the propagator,

$$\begin{aligned}\langle\, z_{\rm f} \,|\, e^{-\beta H} \,|\, z_{\rm i} \,\rangle &= \int \prod_{j=1}^{N-1} \frac{d^2 z_j}{2\pi i}\, \exp\Big(- \mathcal{S}_{\rm E}[\{z_j, \bar z_j\}]/\hbar \Big) \\ \mathcal{S}_{\rm E}[\{z_j, \bar z_j\}]/\hbar &= \sum_{j=1}^{N-1} \Big[\tfrac{1}{2}\, \bar z_j(z_j - z_{j-1}) - \tfrac{1}{2}\, z_j(\bar z_{j+1} - \bar z_j) \Big] + \frac{\epsilon}{\hbar} \sum_{j=1}^{N} H(\bar z_j \,|z_{j-1}) \\ &\qquad + \tfrac{1}{2}\, \bar z_{\rm f}\big(z_{\rm f} - z_{N-1}\big) - \tfrac{1}{2}\, z_{\rm i}\big(\bar z_1 - \bar z_{\rm i}\big) \quad .\end{aligned} \tag{1.274}$$

In the limit $N \to \infty$, we identify the continuum Euclidean action

$$\begin{aligned}\mathcal{S}_{\rm E}[\{z(\tau), \bar z(\tau)\}]/\hbar = \int_0^{\hbar\beta} d\tau \left\{ \frac{1}{2}\left(\bar z\, \frac{dz}{d\tau} - z\, \frac{d\bar z}{d\tau} \right) + \frac{1}{\hbar}\, H(\bar z \,|z) \right\} \\ + \tfrac{1}{2}\, \bar z_{\rm f}\big[z_{\rm f} - z(\hbar\beta) \big] - \tfrac{1}{2}\, z_{\rm i}\big[\bar z(0) - \bar z_{\rm i} \big]\end{aligned} \tag{1.275}$$

and write the continuum expression for the path integral,

$$\langle\, z_{\rm f} \,|\, e^{-\beta H} \,|\, z_{\rm i} \,\rangle = \int\limits_{\substack{z(0)=z_{\rm i} \\ \bar z(\hbar\beta)=\bar z_{\rm f}}} \mathcal{D}[z(\tau), \bar z(\tau)]\, e^{-\mathcal{S}_{\rm E}[\{z(\tau), \bar z(\tau)\}]/\hbar} \quad . \tag{1.276}$$

The corresponding real time expression is given by

$$\langle\, z_{\rm f} \,|\, e^{-iHT/\hbar} \,|\, z_{\rm i} \,\rangle = \int\limits_{\substack{z(0)=z_{\rm i} \\ \bar z(T)=\bar z_{\rm f}}} \mathcal{D}[z(t), \bar z(t)]\, e^{i\mathcal{S}[\{z(t), \bar z(t)\}]/\hbar} \tag{1.277}$$

with

$$\begin{aligned}\mathcal{S}[\{z(t), \bar z(t)\}]/\hbar = \int_0^{T} dt \left\{ \frac{1}{2i}\left(z\, \frac{d\bar z}{dt} - \bar z\, \frac{dz}{dt} \right) - \frac{1}{\hbar}\, H(\bar z \,|z) \right\} \\ + \tfrac{1}{2}\, i\bar z_{\rm f}\big[z_{\rm f} - z(T) \big] - \tfrac{1}{2}\, i z_{\rm i}\big[\bar z(0) - \bar z_{\rm i} \big] \quad .\end{aligned} \tag{1.278}$$

The continuum limit is in a sense justified by examining the discrete equations of motion,

$$\begin{aligned}\frac{1}{\hbar}\frac{\partial \mathcal{S}_{\rm E}}{\partial z_k} &= \bar{z}_k - \bar{z}_{k+1} + \frac{\epsilon}{\hbar}\,\frac{\partial H(\bar{z}_{k+1}|z_k)}{\partial z_k} \\ \frac{1}{\hbar}\frac{\partial \mathcal{S}_{\rm E}}{\partial \bar{z}_k} &= z_k - z_{k-1} + \frac{\epsilon}{\hbar}\,\frac{\partial H(\bar{z}_k|z_{k-1})}{\partial \bar{z}_k} \quad ,\end{aligned} \tag{1.279}$$

which have the sensible continuum limit

$$\hbar\,\frac{d\bar{z}}{d\tau} = \frac{\partial H(\bar{z}\,|z)}{\partial z} \qquad , \qquad \hbar\,\frac{dz}{d\tau} = -\frac{\partial H(\bar{z}\,|z)}{\partial \bar{z}} \tag{1.280}$$

with boundary conditions $\bar{z}(\hbar\beta) = \bar{z}_{\rm f}$ and $z(0) = z_{\rm i}$. Note that there are only two boundary conditions – one on $z(0)$, the other on $\bar{z}(\hbar\beta)$. *The function $z(\tau)$ (or its discrete version z_j) is evolved forward from initial data $z_{\rm i}$, while $\bar{z}(\tau)$ (or $\bar{z}_j$) is evolved backward from final data $\bar{z}_{\rm f}$.* This is the proper number of boundary conditions to place on two first order differential (or finite difference) equations. It is noteworthy that the action of eqn. 1.274 or eqn. 1.275 imposes only a *finite* penalty on *discontinuous* paths.[42] Nevertheless, the paths which extremize the action are continuous throughout the interval $\tau \in (0, \hbar\beta)$. As $z(\tau)$ is integrated forward from $z_{\rm i}$, its final value $z(\hbar\beta)$ will in general be different from $z_{\rm f}$. Similarly, $\bar{z}(\tau)$ integrated backward from $\bar{z}_{\rm f}$ will in general yield an endpoint value $\bar{z}(0)$ which differs from $\bar{z}_{\rm i}$. The differences $z(\hbar\beta) - z_{\rm f}$ and $\bar{z}(0) - \bar{z}_{\rm i}$ are often identified as path discontinuities, but the fact is that the equations of motion know nothing about either $z_{\rm f}$ or $\bar{z}_{\rm i}$. These difference terms do enter in a careful accounting of the action formulae of eqns. 1.274 and 1.275, however.

The importance of the boundary terms is nicely illustrated in a computation of the semiclassical imaginary time propagator for the harmonic oscillator. With $H = \hbar\omega_0\, a^\dagger a$ (dropping the constant term for convenience), we have

$$\begin{aligned}\langle\, z_{\rm f}\,|\,\exp(-\beta\hbar\omega_0\, a^\dagger a)\,|\,z_{\rm i}\,\rangle &= e^{-\frac{1}{2}|z_{\rm f}|^2 - \frac{1}{2}|z_{\rm i}|^2} \sum_{m,n=0}^{\infty} \frac{\bar{z}_{\rm f}^{\,m}\, z_{\rm i}^{\,n}}{\sqrt{m!\,n!}}\,\langle\, m\,|\,\exp(-\beta\hbar\omega_0\, a^\dagger a)\,|\,n\,\rangle \\ &= \exp\left\{ -\tfrac{1}{2}|z_{\rm f}|^2 - \tfrac{1}{2}|z_{\rm i}|^2 + \bar{z}_{\rm f} z_{\rm i}\, e^{-\beta\hbar\omega_0} \right\}\end{aligned} \tag{1.281}$$

The Euclidean action is $L_{\rm E} = \frac{1}{2}\hbar(\bar{z}\dot{z} - z\dot{\bar{z}}) + \hbar\omega_0\,\bar{z}z$, so the equations of motion are

$$\hbar\dot{\bar{z}} = \frac{\partial H}{\partial z} = \hbar\omega_0\,\bar{z} \qquad , \qquad \hbar\dot{z} = -\frac{\partial H}{\partial \bar{z}} = -\hbar\omega_0\, z \tag{1.282}$$

subject to boundary conditions $z(0) = z_{\rm i}$, $\bar{z}(\hbar\beta) = \bar{z}_{\rm f}$. The solution is

$$z(\tau) = z_{\rm i}\, e^{-\omega_0\tau} \qquad , \qquad \bar{z}(\tau) = \bar{z}_{\rm f}\, e^{\omega_0(\tau - \hbar\beta)} \quad . \tag{1.283}$$

[42] In the Feynman path integral, discontinuous paths contribute an infinite amount to the action, and are therefore suppressed.

Along the 'classical path' the Euclidean Lagrangian vanishes: $L_{\rm E} = 0$. The entire contribution to the action therefore comes from the boundary terms:

$$\begin{aligned}\mathcal{S}_{\rm E}^{\rm cl}/\hbar &= 0 + \tfrac{1}{2}\bar{z}_{\rm f}(z_{\rm f} - z_{\rm i}\,e^{-\beta\hbar\omega_0}) - \tfrac{1}{2}z_i(\bar{z}_{\rm f}\,e^{-\beta\hbar\omega_0} - \bar{z}_{\rm i}) \\ &= \tfrac{1}{2}|z_{\rm f}|^2 + \tfrac{1}{2}|z_{\rm i}|^2 - \bar{z}_{\rm f}z_{\rm i}\,e^{-\beta\hbar\omega_0} \quad ,\end{aligned} \tag{1.284}$$

What remains is to compute the fluctuation determinant. We write

$$\begin{aligned}z_j &= z_j^{\rm cl} + \eta_j \\ \bar{z}_j &= \bar{z}_j^{\rm cl} + \bar{\eta}_j\end{aligned} \tag{1.285}$$

and expand the action as

$$\begin{aligned}\mathcal{S}_{\rm E}[\{z_j,\bar{z}_j\}] &= \mathcal{S}_{\rm E}[\{z_j^{\rm cl},\bar{z}_j^{\rm cl}\}] + \frac{\partial^2\mathcal{S}_{\rm E}}{\partial\bar{z}_i\partial z_j}\,\bar{\eta}_i\eta_j + \frac{1}{2}\,\frac{\partial^2\mathcal{S}_{\rm E}}{\partial z_i\partial z_j}\,\eta_i\eta_j + \frac{1}{2}\,\frac{\partial^2\mathcal{S}_{\rm E}}{\partial\bar{z}_i\partial\bar{z}_j}\,\bar{\eta}_i\bar{\eta}_j + \dots \\ &\equiv \mathcal{S}_{\rm E}^{\rm cl} + \frac{\hbar}{2}\begin{pmatrix}\bar{z}_i & z_i\end{pmatrix}\begin{pmatrix}A_{ij} & B_{ij}\\ C_{ij} & A_{ij}^{\rm t}\end{pmatrix}\begin{pmatrix}z_j\\ \bar{z}_j\end{pmatrix} + \dots\end{aligned} \tag{1.286}$$

For general H, we obtain

$$\begin{aligned}A_{ij} &= \delta_{ij} - \delta_{i,j+1} + \frac{\epsilon}{\hbar}\,\frac{\partial^2 H(\bar{z}_i\,|z_j)}{\partial\bar{z}_i\,\partial z_j}\,\delta_{i,j+1} \\ B_{ij} &= \frac{\epsilon}{\hbar}\,\frac{\partial^2 H(\bar{z}_i\,|z_{i-1})}{\partial\bar{z}_i^2}\,\delta_{i,j} \\ C_{ij} &= \frac{\epsilon}{\hbar}\,\frac{\partial^2 H(\bar{z}_{i+1}|z_i)}{\partial z_i^2}\,\delta_{i,j}\end{aligned} \tag{1.287}$$

with i and j running from 1 to $N-1$. The contribution of the fluctuation determinant to the matrix element is then

$$\begin{aligned}\int\prod_{j=1}^{N-1}\frac{d^2\eta_i}{2\pi i}\,\exp\Bigg\{-\frac{1}{2}\begin{pmatrix}\mathsf{Re}\;\eta_k & \mathsf{Im}\;\eta_k\end{pmatrix}\begin{pmatrix}1 & 1\\ -i & i\end{pmatrix}\begin{pmatrix}A_{kl} & B_{kl}\\ C_{kl} & A_{lk}\end{pmatrix}\begin{pmatrix}1 & i\\ 1 & -i\end{pmatrix}\begin{pmatrix}\mathsf{Re}\;\eta_l\\ \mathsf{Im}\;\eta_l\end{pmatrix}\Bigg\} \\ = \mathsf{det}^{-1/2}\begin{pmatrix}A & B\\ C & A^{\rm t}\end{pmatrix}\end{aligned}$$

In the case of the harmonic oscillator discussed above, we have $B_{ij} = C_{ij} = 0$, and since A_{ij} has no elements above its diagonal and $A_{ii} = 1$ for all i, we simply have that the determinant contribution is unity.

1.10 Appendix III : Gauss-Bonnet and Pontrjagin

1.10.1 Gauss-Bonnet theorem

There is a deep result in mathematics, the Gauss-Bonnet theorem, which connects the *local geometry* of a two-dimensional manifold to its *global topology*. The content of the theorem is as follows:

$$\int\limits_{\mathcal{M}} dS\; K = 2\pi\,\chi(\mathcal{M}) = 2\pi \sum_i \mathsf{ind}_{\boldsymbol{x}_i}(\boldsymbol{V})\;, \tag{1.288}$$

where $\mathcal{M}$ is a 2-manifold (a topological space locally homeomorphic to $\mathbb{R}^2$), K is the local *Gaussian curvature* of $\mathcal{M}$, given by $K = (R_1\,R_2)^{-1}$, where $R_{1,2}$ are the principal radii of curvature at a given point, and dS is the differential area element. Here $\boldsymbol{V}(\boldsymbol{x})$ is a vector field on $\mathcal{M}$, and $\mathsf{ind}_{\boldsymbol{x}_i}(\boldsymbol{V})$ refers to the *index* of $\boldsymbol{V}$ at its $i^{\rm th}$ singularity $\boldsymbol{x}_i$. The index is in general defined relative to any closed curve in $\mathcal{M}$, and is given by the winding number of $\boldsymbol{V}(\boldsymbol{x})$ around the curve, *viz.*

$$\mathsf{ind}_{\mathcal{C}}(\boldsymbol{V}) = \oint\limits_{\mathcal{C}} d\boldsymbol{x}\cdot\boldsymbol{\nabla}\tan^{-1}\!\left(\frac{V_2(\boldsymbol{x})}{V_1(\boldsymbol{x})}\right)\quad. \tag{1.289}$$

If $\mathcal{C}$ encloses no singularities, then the index necessarily vanishes, but if $\mathcal{C}$ encloses one or more singularities, the index is an integer, given by the winding number of $\boldsymbol{V}$ around the curve $\mathcal{C}$.

The quantity $\chi(\mathcal{M})$ is called the *Euler characteristic* of $\mathcal{M}$ and is given by $\chi(\mathcal{M}) = 2-2g$, where g is the *genus* of $\mathcal{M}$, which is the number of holes (or handles) of $\mathcal{M}$. Furthermore, $\boldsymbol{V}(\boldsymbol{x})$ can be *any* smooth vector field on $\mathcal{M}$, with $\boldsymbol{x}_i$ the singularity points of that vector field[43].

To apprehend the content of the Gauss-Bonnet theorem, it is helpful to consider an example. Let $\mathcal{M} = \mathbb{S}^2$ be the unit 2-sphere, as depicted in fig. 1.24. At any point on the unit 2-sphere, the radii of curvature are degenerate and both equal to $R = 1$, hence $K = 1$. If we integrate the Gaussian curvature over the sphere, we thus get $4\pi = 2\pi\,\chi\big(\mathbb{S}^2\big)$, which says $\chi(\mathbb{S}^2) = 2-2g = 2$, which agrees with $g = 0$ for the sphere. Furthermore, the Gauss-Bonnet theorem says that *any* smooth vector field on $\mathbb{S}^2$ *must* have a singularity or singularities, with the total index summed over the singularities equal to $+2$. The vector field sketched in the left panel of fig. 1.24 has two index $+1$ singularities, which could be taken at the north and south poles, but which could be anywhere. Another possibility, depicted in the right panel of fig. 1.24, is that there is a one singularity with index $+2$.

In fig. 1.25 we show examples of manifolds with genii $g = 1$ and $g = 2$. The case $g = 1$ is the familiar 2-torus, which is topologically equivalent to a product of circles: $\mathbb{T}^2 \cong \mathbb{S}^1 \times \mathbb{S}^1$, and is thus coordinatized by two angles θ_1 and θ_2. A smooth vector field pointing in the direction of increasing θ_1 never vanishes, and thus has no singularities, consistent with $g = 1$ and $\chi\big(\mathbb{T}^2\big) = 0$.

[43] The singularities $\boldsymbol{x}_i$ are fixed points of the dynamical system $\dot{\boldsymbol{x}} = \boldsymbol{V}(\boldsymbol{x})$.

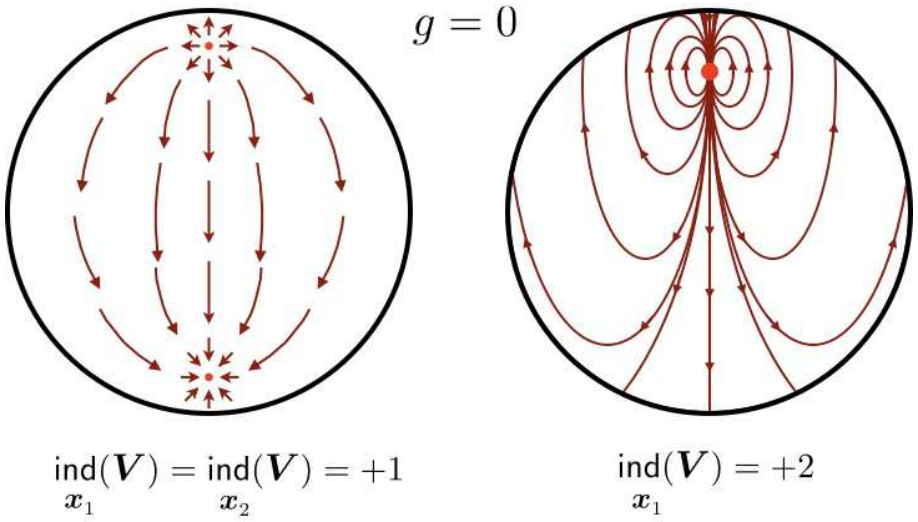


Figure 1.24: Two smooth vector fields on the sphere $\mathbb{S}^2$, which has genus $g = 0$. Left panel: two index $+1$ singularities. Right panel: one index $+2$ singularity.

Topologically, one can define a torus as the quotient space $\mathbb{R}^2/\mathbb{Z}^2$, or as a square with opposite sides identified. This is what mathematicians call a 'flat torus' – one with curvature $K = 0$ everywhere. Of course, such a torus cannot be embedded in three-dimensional Euclidean space; a two-dimensional figure embedded in a three-dimensional Euclidean space inherits a metric due to the embedding, and for a physical torus, like the surface of a bagel, the Gaussian curvature is only zero *on average*.

The $g = 2$ surface $\mathcal{M}$ shown in the right panel of fig. 1.25 has Euler characteristic $\chi(\mathcal{M}) = -2$, which means that any smooth vector field on $\mathcal{M}$ must have singularities with indices totalling -2. One possibility, depicted in the figure, is to have two saddle points with index -1; one of these singularities is shown in the figure (the other would be on the opposite side).

1.10.2 The Pontrjagin index

Consider an N-dimensional vector field $\dot{\boldsymbol{x}} = \boldsymbol{V}(\boldsymbol{x})$, and let $\hat{\boldsymbol{n}}(\boldsymbol{x})$ be the unit vector field defined by $\hat{\boldsymbol{n}}(\boldsymbol{x}) = \boldsymbol{V}(\boldsymbol{x})/|\boldsymbol{V}(\boldsymbol{x})|$. Consider now a unit sphere in $\hat{\boldsymbol{n}}$ space, which is of dimension $(N-1)$. If we integrate over this surface, we obtain

$$\Omega_N = \oint d\sigma_a \, n^a = \frac{2\pi^{(N-1)/2}}{\Gamma\left(\frac{N-1}{2}\right)} \,, \tag{1.290}$$

which is the surface area of the unit sphere $\mathbb{S}^{N-1}$. Thus, $\Omega_2 = 2\pi$, $\Omega_3 = 4\pi$, $\Omega_4 = 2\pi^2$, *etc.*

Now consider a change of variables to those over the surface of the sphere, $(\xi_1, \ldots, \xi_{N-1})$. We

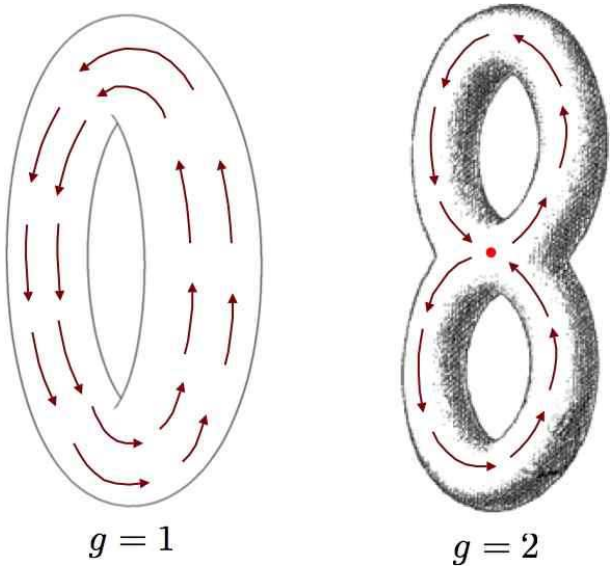


Figure 1.25: Smooth vector fields on the torus $\mathbb{T}^2$, and on a 2-manifold $\mathcal{M}$ of genus $g = 2$

then have

$$\Omega_N = \oint\limits_{\mathbb{S}^{N-1}} d\sigma_a \, n^a = \oint d^{N-1}\xi \, \epsilon_{a_1 \cdots a_N} \, n^{a_1} \, \frac{\partial n^{a_2}}{\partial \xi_1} \cdots \frac{\partial n^{a_N}}{\partial \xi_{N-1}} \tag{1.291}$$

The topological charge is then

$$Q = \frac{1}{\Omega_N} \oint d^{N-1}\xi \, \epsilon_{a_1 \cdots a_N} \, n^{a_1} \, \frac{\partial n^{a_2}}{\partial \xi_1} \cdots \frac{\partial n^{a_N}}{\partial \xi_{N-1}} \tag{1.292}$$

The quantity Q is an *integer topological invariant* which characterizes the map from the surface $(\xi_1, \ldots, \xi_{N-1})$ to the unit sphere $|\hat{\boldsymbol{n}}| = 1$. In mathematical parlance, Q is known as the *Pontrjagin index* of this map.

This analytical development recapitulates some basic topology. Let $\mathcal{M}$ be a topological space and consider a map from the circle $\mathbb{S}^1$ to $\mathcal{M}$. We can compose two such maps by merging the two circles, as shown in fig. 1.26. Two maps are said to be *homotopic* if they can be smoothly deformed into each other. Any two homotopic maps are said to belong to the same *equivalence class* or *homotopy class*. For general $\mathcal{M}$, the homotopy classes may be multiplied using the composition law, resulting in a group structure. The group is called the *fundamental group* of the manifold $\mathcal{M}$, and is abbreviated $\pi_1(\mathcal{M})$. If $\mathcal{M} = \mathbb{S}^2$, then any such map can be smoothly contracted to a point on the 2-sphere, which is to say a trivial map. We then have $\pi_1(\mathcal{M}) = 0$. If $\mathcal{M} = \mathbb{S}^1$, the maps can wind nontrivially, and the homotopy classes are labeled by a single integer winding number: $\pi_1(\mathbb{S}^1) = \mathbb{Z}$. The winding number of the composition of two such maps is the sum of their individual winding numbers. If $\mathcal{M} = \mathbb{T}^2$, the maps can wind nontrivially around either of the two cycles of the 2-torus. We then have $\pi_1(\mathbb{T}^2) = \mathbb{Z}^2$, and in general $\pi_1(\mathbb{T}^n) = \mathbb{Z}^n$. This makes good sense, since an n-torus is topologically equivalent to

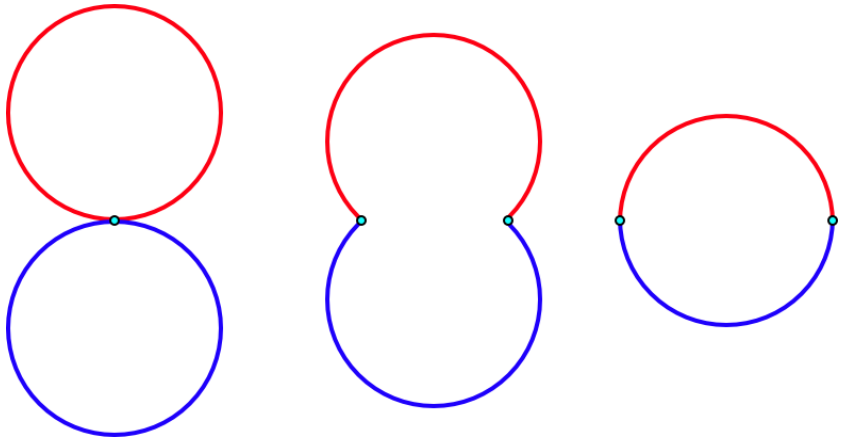


Figure 1.26: Composition of two circles. The same general construction applies to the merging of n-spheres $\mathbb{S}^n$, called the *wedge sum*.

a product of n circles. In some cases, $\pi_1(\mathcal{M})$ can be nonabelian, as is the case when $\mathcal{M}$ is the genus $g = 2$ structure shown in the right hand panel of fig. 1.25.

In general we define the n^{th} *homotopy group* $\pi_n(\mathcal{M})$ as the group under composition of maps from $\mathbb{S}^n$ to $\mathcal{M}$. For $n \geq 2$, $\pi_n(\mathcal{M})$ is abelian. If $\dim(\mathcal{M}) < n$, then $\pi_n(\mathcal{M}) = 0$. In general, $\pi_n(\mathbb{S}^n) = \mathbb{Z}$. These n^{th} homotopy classes of the n-sphere are labeled by their associated Pontrjagin index Q.

Chapter 2

Integer Quantum Hall Effect

2.1 Continuum Percolation

2.1.1 Dynamics in the LLL

Recall the classical equation of motion for an electron in a field $\boldsymbol{B}$ subject to a potential $V(\boldsymbol{r})$,

$$m\ddot{\boldsymbol{r}} = -\boldsymbol{\nabla} V - \frac{e}{c}\,\dot{\boldsymbol{r}} \times \boldsymbol{B} \quad . \tag{2.1}$$

Averaging over the fast classical cyclotron motion $\boldsymbol{\xi}(t)$, we obtained the dynamics of the guiding-center $\boldsymbol{\mathcal{R}}(t)$,

$$\frac{d\boldsymbol{\mathcal{R}}}{dt} = \frac{\ell^2}{\hbar}\,\hat{\boldsymbol{z}} \times \boldsymbol{\nabla} V_{\rm eff}(\boldsymbol{\mathcal{R}}) \quad , \tag{2.2}$$

where the effective potential is $V_{\rm eff}(\boldsymbol{\mathcal{R}}) = V(\boldsymbol{\mathcal{R}}) + \frac{1}{2}\langle \boldsymbol{\xi}^2 \rangle\, \nabla^2 V(\boldsymbol{\mathcal{R}}) + \ldots$, and where we have taken $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$. Thus, $dV_{\rm eff}\big(\boldsymbol{\mathcal{R}}(t)\big)/dt = 0$ and the guiding-center moves along an equipotential.

At the quantum level, recall how in chapter 1 we derived the LLL-projected potential,

$$\widetilde{V}(\boldsymbol{\mathcal{R}}) = \langle\, 0 \,|\, V \,|\, 0 \,\rangle = \int \frac{d^2k}{(2\pi)^2}\, \hat{V}(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{\mathcal{R}}}\, e^{-\boldsymbol{k}^2\ell^2/4} = \big(1 + \tfrac{1}{4}\,\ell^2\,\nabla^2 + \ldots\big) V(\boldsymbol{\mathcal{R}}) \quad , \tag{2.3}$$

where $\boldsymbol{\mathcal{R}}$ is the guiding-center position operator, the Cartesian components of which satisfy $[\mathcal{X}, \mathcal{Y}] = -i\ell^2$. Thus we have the equivalences

$$\mathcal{X} = \frac{\ell^2}{i}\,\frac{\partial}{\partial \mathcal{Y}} \qquad , \qquad \mathcal{Y} = i\ell^2\,\frac{\partial}{\partial \mathcal{X}} \quad . \tag{2.4}$$

Thus the Ehrenfest equations of motion are

$$\frac{d\langle \mathcal{X} \rangle}{dt} = -\frac{\ell^2}{\hbar}\,\Big\langle\, \frac{\partial \widetilde{V}}{\partial \mathcal{Y}} \,\Big\rangle \qquad , \qquad \frac{d\langle \mathcal{Y} \rangle}{dt} = +\frac{\ell^2}{\hbar}\,\Big\langle\, \frac{\partial \widetilde{V}}{\partial \mathcal{X}} \,\Big\rangle \quad . \tag{2.5}$$

At the semiclassical level, we remove the brackets, replace $\langle \mathcal{X} \rangle \to X$ and $\langle \mathcal{Y} \rangle \to Y$, and write $\dot{X} = -\frac{\ell^2}{\hbar}\frac{\partial \widetilde{V}}{\partial Y}$ and $\dot{Y} = +\frac{\ell^2}{\hbar}\frac{\partial \widetilde{V}}{\partial X}$. We can reproduce these results from the coherent state path integral approach. Recall that the complexified guiding-center position operator is $\mathcal{R} = \sqrt{2}\,\ell b^\dagger$, hence the coherent state path integral action is given by Eqn. 1.278, replacing $z = \bar{R}/\sqrt{2}\,\ell$, *i.e.*

$$\mathcal{S}\big[\{R(t), \bar{R}(t)\}\big]/\hbar = \int_0^T dt \left\{ \frac{1}{4i\ell^2}\left(\bar{R}\,\frac{dR}{dt} - R\,\frac{d\bar{R}}{dt}\right) - \frac{1}{\hbar}\,\widetilde{V}(R\,|\bar{R})\right\} + \Delta\mathcal{S}/\hbar \quad , \tag{2.6}$$

where

$$\Delta\mathcal{S} = \frac{i\hbar}{4\ell^2}\Big(R_{\rm f}\big[\bar{R}_{\rm f} - \bar{R}(T)\big] - \bar{R}_{\rm i}\big[R(0) - R_{\rm i}\big]\Big) \quad . \tag{2.7}$$

is the boundary discontinuity term, which does not affect the equations of motion. Taking the functional variation with respect to R and $\bar{R}$ yields the complexified equations of motion,

$$\frac{dR}{dt} = \frac{2i\ell^2}{\hbar}\,\frac{\partial \widetilde{V}(R\,|\bar{R})}{\partial \bar{R}} \quad , \qquad \frac{d\bar{R}}{dt} = -\frac{2i\ell^2}{\hbar}\,\frac{\partial \widetilde{V}(R\,|\bar{R})}{\partial R} \quad , \tag{2.8}$$

which are indeed the complexified forms of

$$\dot{X} = -\frac{\ell^2}{\hbar}\,\frac{\partial \widetilde{V}}{\partial Y} \quad , \qquad \dot{Y} = +\frac{\ell^2}{\hbar}\,\frac{\partial \widetilde{V}}{\partial X} \quad . \tag{2.9}$$

In vectorized form, $\dot{\boldsymbol{R}} = \hbar^{-1}\ell^2\,\hat{\boldsymbol{z}} \times \boldsymbol{\nabla}\widetilde{V}$.

2.1.2 Electrons in a smooth random potential

In heterojunction inversion layers, the potential $\widetilde{V}(\boldsymbol{R})$ arises from the electrical potential due to the recessed donor ions (Si^+ substituting for Al in $Al_xGa_{1-x}As$). While the displacement of the Si^+ dopant ions in the direction ($\hat{\boldsymbol{z}}$) perpendicular to the inversion layer is rather precisely controlled ("δ-doping"), they are located pretty much randomly in the (x, y) plane, hence the potential may be taken to be random, though smoothed out on a scale on the order of the distance between the GaAs-$Al_xGa_{1-x}As$ interface and the dopant layer, which is typically several hundreds of Ångstroms. Landau level projection further smooths the potential by further suppressing high spatial frequencies via the $\exp(-\boldsymbol{k}^2\ell^2/4)$ factor. A mock-up of such a smooth random potential is shown in Fig. 2.1. One can generate such $\widetilde{V}\boldsymbol{R})$ from a distribution functional

$$P\big[\widetilde{V}(\boldsymbol{R})\big] = P[0]\,\exp\left\{-\frac{1}{2\gamma}\int d^2R\,\Big[\widetilde{V}^2 + \lambda^2\big(\boldsymbol{\nabla}\widetilde{V}\big)^2\Big]\right\} \quad , \tag{2.10}$$

where $P[0]$ ensures normalization of the functional integral $\int D\widetilde{V}\,P\big[\widetilde{V}\big] = 1$. Here, λ is the length scale over which $\widetilde{V}(\boldsymbol{R})$ is correlated. Indeed, for the above distribution functional, the

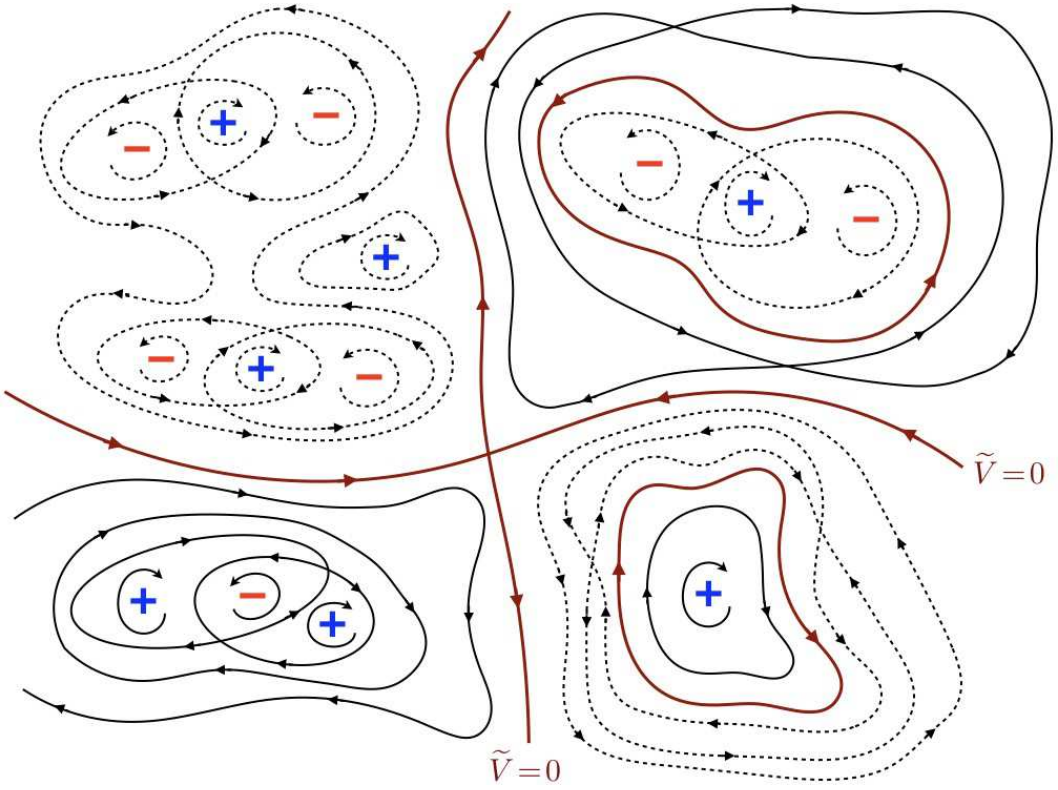


Figure 2.1: Contour plots for a symmetrically random potential $\widetilde{V}(\boldsymbol{R})$. Electrons move clockwise around peaks (+) and counterclockwise around valleys (−). Solid lines indicate level sets with $\widetilde{V}(\boldsymbol{R}) > 0$; dashed lines correspond to $\widetilde{V}(\boldsymbol{R}) < 0$. The brown curves lie at $\widetilde{V}(\boldsymbol{R}) = 0$, which is the continuum percolation threshold.

correlation function is of the two-dimensional Ornstein-Zernike form,

$$\left\langle \widetilde{V}(\boldsymbol{R})\,\widetilde{V}(\boldsymbol{0})\right\rangle = \frac{\gamma}{2\pi\lambda^2}\,K_0(R/\lambda) \quad , \tag{2.11}$$

where $K_0(z)$ is the Hankel function of imaginary argument[1], whose asymptotic behavior is

$$K_0(z) = \begin{cases} -\ln z + \ln 2 - \mathsf{C} + \mathcal{O}(z^2 \ln z) & z \to 0 \\ (\pi/2z)^{1/2}\exp(-z) \times \left\{1 + \mathcal{O}(z^{-1})\right\} & z \to \infty \end{cases} \quad , \tag{2.12}$$

where $\mathsf{C} = 0.57721\ldots$ is the Euler-Mascheroni constant. The divergence as $R \to 0$ can be cured by imposing a cutoff at high spatial frequency $k = \Lambda$, rendering $\left\langle \widetilde{V}^2(\boldsymbol{0})\right\rangle$ finite.

The equipotentials (*i.e.* the level sets) of a given random $\widetilde{V}(\boldsymbol{R})$ will appear as in Fig. 2.1, with peaks, valleys, and saddle points. The LLL dynamics in a field $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$ are such that electrons circle clockwise around peaks and counterclockwise around valleys, as depicted in the figure.

[1]See Gradshteyn and Ryzhik §8.4.

For symmetrically distributed $\widetilde{V}(\boldsymbol{R})$, there will be a unique value $\widetilde{V}(\boldsymbol{R}) = 0$ where the level set is of infinite continuous extent. Electrons with energies $\widetilde{V} = 0$ can percolate across the entire sample, which is an infinite distance in the thermodynamic limit.

Semiclassically, the LLL electron wavefunctions are localized along the equipotentials. Recall the wavefunctions in the symmetric gauge and in the absence of a potential are given by $\psi_m(\boldsymbol{r}) = C_m\, z^m \exp(-|z|^2/4\ell^2)$. Maximizing $\psi_m(\boldsymbol{r})|^2$, one finds $r_m^2 = 2m\ell^2$, which is a ring enclosing an area $A_m = 2\pi m\ell^2$. Thus, increasing m by $\Delta m = 1$ is associated with a concomitant increase in area by the quantum $2\pi\ell^2$. The semiclassical wavefunctions obey the same quantization rule $A_m = 2\pi m\ell^2$, except they are localized along equipotentials of $\widetilde{V}(\boldsymbol{r})$ rather than along circles. If we parameterize an equipotential curve $\widetilde{V}(\boldsymbol{r}) = E_{\rm F}$ by a distance u along the curve and a distance v locally perpendicular to it, then the semiclassical eigenfunctions, following Trugman[2], take the form

$$\psi(u,v) = |\boldsymbol{\nabla}\widetilde{V}(u,0)|^{-1/2}\, H_n(v/\ell)\, \exp(-v^2/2\ell^2)\, e^{i\chi(u,v)} \quad , \tag{2.13}$$

where $\chi(u,v)$ is a gauge-dependent phase function whose winding around the equipotential increases by 2π with each consecutive semiclassical energy eigenstate. Here we have included the LL index n; note that this is essentially the Landau strip wavefunction written in local coordinates (u,v). It is valid provided $\ell \ll b$ where b is the local radius of curvature of the contour, and if $|\boldsymbol{\nabla}\widetilde{V}| \ll \hbar\omega_{\rm c}/\ell$. Because electrons are fermions, these semiclassical levels will be filled up to the Fermi energy. The contour at $\widetilde{V}(\boldsymbol{r}) = E_{\rm F}$ represents the highest occupied electronic energy level. The following vivid analogy may be helpful:

> *Imagine $\widetilde{V}(\boldsymbol{r})$ is the height function of a random landscape. After a long period of rain, every part of the landscape for which $\widetilde{V}(\boldsymbol{r}) < E_{\rm F}$ is under water. You are constrained to walk in such a way that your left foot is always under water, and your right foot is always on dry land.*

When the Fermi level $E_{\rm F}$ is low, you are constrained to walk in a counterclockwise direction (as viewed from above) about isolated puddles. When the Fermi level $E_{\rm F}$ is high, you are constrained to walk in a clockwise direction around isolated islands. In each case, you don't get very far from wherever you started. Now you know what a classical electron in a random potential and a large magnetic field feels like.

For a given realization $\widetilde{V}(\boldsymbol{r})$ of the random potential, the density of level sets, per unit energy, is given by

$$\rho(E) = \frac{1}{A}\int d^2r\, \delta\big(E - \widetilde{V}(\boldsymbol{r})\big) \quad , \tag{2.14}$$

where A is the total area. The electronic density of states[3] is $g(E) = N_\phi^{-1}\sum_m \delta(E - E_m)$, where m is an eigenstate label within the LLL. In the limit $B \to \infty$, we have $g(E) = \rho(E)$, which is

[2] See S. A. Trugman, *Phys. Rev. B* **27**, 7539 (1983).

[3] Here $g(E)$ is defined to be the DOS per unit energy per unit flux.

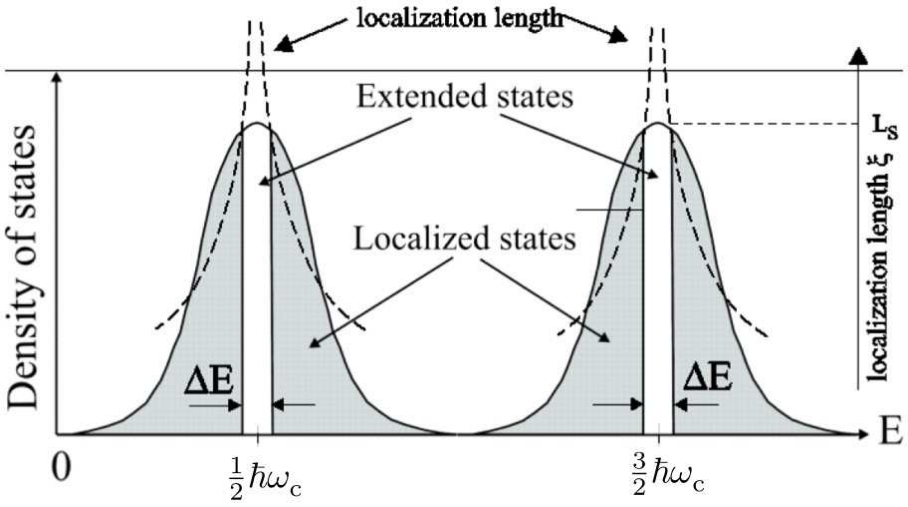


Figure 2.2: Density of states, correlation length, and mobility gaps in disorder-broadened Landau levels. In the thermodynamic limit, extended states exist only at unique energies corresponding to the (quantum) continuum percolation threshold at the centers of each Landau level. For systems of finite linear dimension L, a range of states with $\xi(E) > L$ may be considered as extended. *Image: J. Oswald, DOI:10.5772/62926.*

to say that each eigenstate is associated with a quantum of area $2\pi\ell^2$. Note the normalization condition $\int_{-\infty}^{\infty} dE\, \rho(E) = \int_{-\infty}^{\infty} dE\, g(E) = 1$. We are interested in the properties of the eigenstates as a function of their energy E. In particular, what is their typical spatial extent? Let the standard deviation of the random potential be $\Delta = \left\langle \widetilde{V}^2(\mathbf{0}) \right\rangle^{1/2}$. For E/Δ sufficiently negative, only the lowest valleys will support occupied electronic states. Similarly, for E/Δ sufficiently positive, only the highest peaks will support *unoccupied* electronic states. As $|E|$ decreases, the spatial extent of the equipotentials $\widetilde{V}(\boldsymbol{r}) = E$ increases. If $\widetilde{V}(\boldsymbol{r})$ is symmetrically distributed, then there will be a unique critical energy $E_c = 0$ at which the typical size of the equipotentials diverges, as $\xi(E) \propto |E|^{-\nu}$, where $\nu = \frac{4}{3}$ is the correlation length exponent for two-dimensional *percolation*.

2.1.3 Percolation theory

Percolation is a geometric critical phenomenon describing the clustering and the emergence of an infinite connected network in random systems[4]. We first describe the setting for *site percolation*. Consider a lattice in which each site is randomly occupied with probability $p \in [0, 1]$. One defines a *cluster* as a maximal connected set of occupied sites[5]. An *s-cluster* is defined to be a cluster of size s. The probability that a given site belongs to a cluster of *infinite* extent is

[4]See, *e.g.*, D. Stauffer, *Phys. Reg.* **54**, 1 (1979) and J. W. Essam, *Rep. Prog. Phys.* **43**, 53 (1980).
[5]Maximal in the sense that all occupied sites connected to the cluster are accounted to be in the cluster.

called the *percolation probability*, $P_\infty \equiv P(p)$. The *percolation threshold* is the largest value of p for which $P(p) = 0$. For $p < p_{\rm c}$ one has $P(p) = 0$, but for $p - p_{\rm c}$ small and positive, $P(p) \propto (p - p_{\rm c})^\beta$, where β is a critical exponent. Just as in a magnetic system, where the order parameter is the magnetization $\boldsymbol{M}(T) \propto (T_{\rm c} - T)_+^\beta$, in percolation theory the order parameter is $P(p)$[6].

Let $n_s(p)$ be the number of s-clusters per lattice site. Then for each lattice site there are three possibilities: (i) the site may be unoccupied, with probability $1-p$, (ii) the site may be occupied and a member of a finite cluster of size s, with probability n_s, or (iii) the site may be occupied and a member of an infinite cluster, with probability $p\,P(p)$. Thus,

$$(1-p) + \sum_{s=1}^{\infty} s\,n_s(p) + p\,P(p) = 1 \quad . \tag{2.15}$$

Note that this entails $\sum_{s=1}^\infty s\,n_s(p) = p\big(1 - P(p)\big)$. Examples of percolation clusters on are depicted in Fig. 2.3 for the square lattice[7]. As an application, consider a dilute Ising magnet at temperatures $T \ll J/k_{\rm B}$, where J is the exchange energy. The magnetization $M(T, H, p)$ is given by

$$M(T, H, p) = \pm P(p) + p^{-1} \sum_{s=1}^{\infty} s\,n_s(p)\,\tanh(s\mu H/k_{\rm B}T) \quad . \tag{2.16}$$

For $p < p_{\rm c}$, only finite clusters are present, and there is zero magnetization at $H = 0$. For $p > p_{\rm c}$, there is an infinite cluster, which immediately polarizes for any finite H. In thermodynamic equilibrium, we have $\pm P(p) = P(p)\,\mathrm{sgn}(H)$, but it may be that the infinite cluster gets stuck in a metastable state, *i.e.* that there is hysteresis. Another application of percolation theory is to the properties of random mixtures of conducting and nonconducting elements, such as aluminum and glass marbles, or random resistor networks[8].

In the vicinity of $p = p_{\rm c}$, the following critical properties pertain:

$$\begin{aligned} \sum_{s=1}^{\infty} n_s(p) &\propto |p - p_{\rm c}|^{2-\alpha} + \text{nst} \quad , \quad \sum_{s=1}^{\infty} s\,n_s(p) \propto (p - p_{\rm c})_+^\beta + \text{nst} \\ \sum_{s=1}^{\infty} s^2\,n_s(p) &\propto |p - p_{\rm c}|^{-\gamma} + \text{nst} \quad , \quad \sum_{s=1}^{\infty} s\,n_s(p_{\rm c})\,e^{-hs} \propto h^{1/\delta} + \text{nst} \quad , \end{aligned} \tag{2.17}$$

with the 'field' h small and positive, and where "nst" means "non-singular terms". One also defines the *correlation length* $\xi(p)$ to be the typical diameter of finite clusters. Let $g(\boldsymbol{r}, p)$ be

[6] We define $x_+ \equiv x\,\Theta(x)$.

[7] One might think that if the set of occupied sites does not percolate, *i.e.* if $p < p_{\rm c}$, that the unoccupied sites must percolate, but this is false. Clearly on any bipartite lattice if all the A sublattice sites are occupied and all the B sublattice sites are unoccupied, then *neither* the occupied nor the unoccupied sites percolates. Indeed in this example the only clusters are of size $s = 1$.

[8] See, *e.g.* S. Kirkpatrick, *Rev. Mod. Phys.* **45**, 574 (1973).

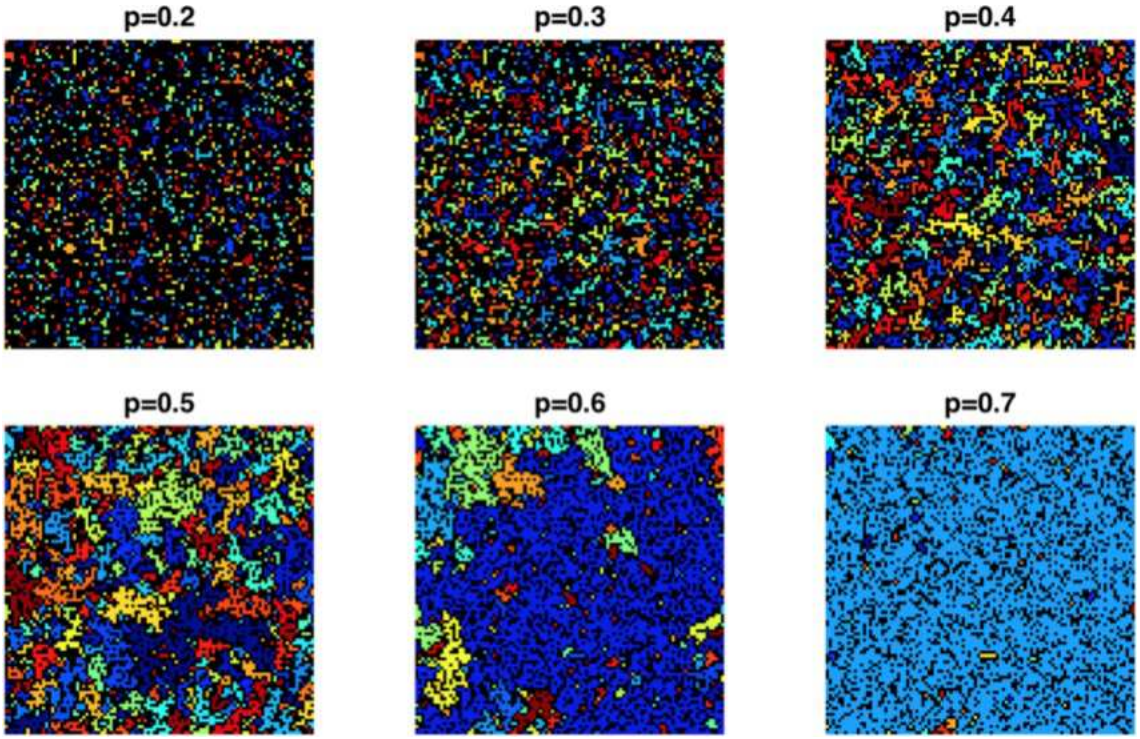


Figure 2.3: Site percolation clusters on the square lattice ($p_c \simeq 0.5927$). Each distinct cluster appears as a different color. Unoccupied sites are shown in black. The infinite cluster for $p = 0.7$ is shown in light blue. From A. Malthe-Sorenssen, *Percolation and Disordered Systems – A Numerical Approach* (unpublished, 2015).

the *pair connectivity function*, defined as the probability that two occupied sites separated by a distance $\boldsymbol{r}$ belong to the same finite cluster. One then has

$$\xi^2(p) = \frac{\sum_{\boldsymbol{r}} \boldsymbol{r}^2 g(\boldsymbol{r}, p)}{\sum_{\boldsymbol{r}} g(\boldsymbol{r}, p)} \quad . \tag{2.18}$$

The scaling hypothesis

Following Stauffer, in the vicinity of $p = p_c$, we adopt a scaling *Ansatz* which says that the critical behavior is dominated by clusters of size $s_\xi \propto |p - p_c|^{-1/\sigma}$, where σ is a universal critical exponent. Precisely at $p = p_c$, the singular part of the cluster size distribution is presumed to behave as $n(p_c) \propto s^{-\tau}$, where τ is another universal critical exponent. Thus, the scaling hypothesis entails the relation $n_s(p) = n_s(p_c)\, \phi(s/s_\xi)$ with $\phi(0) = 1$, *i.e.*

$$n_s(p) \propto s^{-\tau} \phi_\pm\big(|p - p_c|^{1/\sigma} s\big) \quad , \tag{2.19}$$

where the $\pm$ sign is for $p \gtrless p_c$. It is important to understand that s_ξ is the cluster size which dominates in the *singular part* of $\sum_{s=1}^{\infty} s\, n_s(p)$; the smooth contributions are dominated by much

d	α	β	γ	δ	η	ν	σ	τ	$d_{\rm f}$
2	$-2/3$	$5/36$	$43/18$	$91/5$	$5/24$	$4/3$	$36/91$	$187/91$	$91/48$
3	$-0.625(3)$	$0.418(1)$	$1.805(2)$	$5.29(6)$	$-0.059(9)$	$0.875(1)$	$0.445(1)$	$2.190(2)$	$2.530(1)$
4	$-0.756(4)$	$0.657(9)$	$1.435(1)$	$3.198(6)$	$-0.0929(9)$	$0.689(1)$	$0.476(5)$	$2.313(3)$	$3.056(7)$

Table 2.1: Critical exponents for percolation. η is the anomalous exponent describing the power law decay of correlations at criticality, and $d_{\rm f}$ is the fractal dimension of the percolation cluster. For $d > 2$, there is some variation reported in numerical computations of the critical exponents. *Source: Wikipedia (Percolation Critical Exponents).*

smaller clusters. If we further define $g_{s,t}$ to be the number of distinct cluster configurations with total size s and perimeter t, then the average number of s-clusters with perimeter t is given by $n_{s,t}(p) = g_{s,t}\, p^s\, (1-p)^t$. An analog for the partition function can be defined for the percolation problem, *viz.*[9]

$$Z(p,h) = \sum_{s=1}^{\infty}\sum_{t=1}^{\infty} g_{s,t}\, p^s\, q^t\, e^{-ht} = \sum_{s=1}^{\infty}\sum_{t=1}^{\infty} n_{s,t}(p)\, e^{-ht} \quad , \tag{2.20}$$

where $q = 1-p$. Writing $Z_s(p,h) = \sum_t g_{s,t}\, q^t\, e^{-ht}$, one has

$$t_s(p) = \frac{\sum_t t\, n_{s,t}(p)}{\sum_t n_{s,t}(p)} = \frac{\partial \ln Z_s(p, h=0)}{\partial \ln q} = \frac{qs}{p} - q\,\frac{\partial \ln n_s(p)}{\partial p} \quad . \tag{2.21}$$

For large s, the second term is known to behave as s^ζ with $\zeta < 1$, hence in the large s limit we have $t_s = (p^{-1}-1)\, s \propto s$. Thus, the large clusters are highly ramified, with $t_s \sim s$. Note that summing $n_{s,t}(p)$ over the perimeter t gives $n_s(p) = \sum_t n_{s,t}(p)$.

From the scaling relations, we may obtain all the critical exponents in terms of σ and τ. For example,

$$\sum_{s=1}^{\infty} s\, n_s(p) \sim \int_1^{\infty} ds\; s^{1-\tau}\, \phi_\pm\big(|\delta p|^{1/\sigma} s\big) \sim \int_{|\delta p|^{1/\sigma}}^{\infty} du\; e^{1-\tau}\, \phi(u)\cdot |\delta p|^{(\tau-2)/\sigma} \propto (\delta p)_+^{\beta} \tag{2.22}$$

where $\delta p \equiv p - p_{\rm c}$. Thus we conclude $\beta = (\tau-2)/\sigma$. Similarly,

$$\sum_{s=1}^{\infty} s\, n_s(p_{\rm c}) e^{-hs} \sim \int_1^{\infty} ds\; s^{1-\tau}\, \phi(0)\, e^{-hs} \sim \phi(0) \int_h^{\infty} du\; u^{1-\tau}\, e^{-u}\cdot h^{2-\tau} \propto h^{1/\delta} \tag{2.23}$$

whence $\delta = 1/(\tau-2)$. The full set of exponents is given by

$$\alpha = 2 + \frac{1-\tau}{\sigma} \quad , \quad \beta = \frac{\tau-2}{\sigma} \quad , \quad \gamma = \frac{3-\tau}{\sigma} \quad , \quad \delta = \frac{1}{\tau-2} \quad . \tag{2.24}$$

[9] Note that the minimum value t can take for nonzero $g_{s,t}$ is $t = z$, the lattice coordination number.

Finally, assuming hyperscaling, which is to say that the singular part of the free energy density scales as $\big[\xi(p)\big]^{-d}$, one may derive the correlation length exponent $\nu = (\tau-1)/\sigma d$. Values for the critical exponents is $d = 2, 3, 4$ are listed in Tab. 2.1. For reference, the anomalous correlation exponent η, which governs $g(\boldsymbol{r}, p_{\rm c}) \propto r^{-d+2-\eta}$, is given by $\eta = 2 + d - \frac{2d}{\tau-1}$.

There is another type of percolation, called *bond percolation*, in which the links of the lattice are occupied (open) with probability p and vacant (closed) with probability $1-p$. In bond percolation, a cluster is defined to be a maximal group of connected bonds. While the percolation thresholds $p_{\rm c}$ on a given lattice in general differ for site and bond percolation, the critical exponents, being universal, do not. Values of $p_{\rm c}$ for site and bond[10] percolation on some common lattices are given in Tab. 2.2.

It is interesting to note[11] that while the critical probability $p_{\rm c}$ for site percolation varies significantly from lattice to lattice, even holding the dimensionality d fixed, when one accounts for the lattice *filling factor* f, defined to be the fraction of the total volume filled when the lattice points are surrounded by hard spheres of maximal radius[12], the product $\phi_{\rm c} \equiv f p_{\rm c}$ is approximately independent of the lattice type and depends only on dimensionality[13], with $\phi_{\rm c}(d=2) \simeq 0.44$ and $\phi_{\rm c}(d=3) \simeq 0.15$. These values also approximately hold for random networks. Thus, if you fill a volume randomly with glass and aluminum marbles, the onset of bulk conduction will occur when the total volume fraction of aluminum exceeds about 15%.

2.1.4 Continuum percolation

Let's now return to our original problem of characterizing the level sets of a smooth random potential $\widetilde{V}(\boldsymbol{r})$. We can associate with this problem a correlated site percolation problem, where the site occupation probability p is given by the fraction of the landscape which lies 'under water', *i.e.* with $\widetilde{V}(\boldsymbol{r}) \le E$, is given by

$$p(E) = \int_{-\infty}^{E} dE'\, \rho(E') \quad . \tag{2.25}$$

As mentioned above, for the continuum percolation problem, the critical energy is $E_{\rm c} = 0$, and therefore $p(0) = p_{\rm c}$. Within the classical continuum percolation picture, the typical cluster size grows as $\xi(E) \propto |E|^{-4/3}$ for $E \approx 0$.

In any physical setting,, the sample dimensions will be finite, and the 2DEG is confined to a

[10] Interesting factoid: $p_{\rm c}^{\rm bond}(d=2) \simeq 2/z$ while $p_{\rm c}^{\rm bond}(d=3) \simeq 3/2z$, where z is the lattice coordination number.
[11] H. Scher and R. Zallen, *J. Chem. Phys.* **53**, 3749 (1970).
[12] *I.e.*, spheres centered on two neighboring lattice sites are tangent.
[13] We stress that the independence of $\phi_{\rm c}$ on lattice structure is not rigorously true, but only approximately so.

d	lattice	z	p_c^{site}	p_c^{bond}
1	chain	2	$1.000\ldots$	$1.000\ldots$
2	honeycomb	3	0.6962	$0.65270\ldots$
2	kagome	4	$0.65260\ldots$	0.5244
2	square	4	0.592746	$0.500\ldots$
2	triangular	6	$0.500\ldots$	$0.34729\ldots$
3	diamond	4	0.43	0.388
3	simple cubic	6	0.3116	0.2488
3	bcc	8	0.246	0.1803
3	fcc	12	0.1998	0.119
4	hypercubic	8	0.197	0.1601
∞	Bethe lattice	z	$1/(z-1)$	$1/(z-1)$

Table 2.2: Site and bond percolation thresholds on various lattices. *Source: Wikipedia (Percolation Threshold)*.

Hall bar of roughly rectangular shape[14]. At the edges of the Hall bar, then, there is a confining potential which rises to some high value. This keeps the electrons from spilling out into the vacuum. The situation is schematically illustrated in Fig. 2.4. When the Fermi energy $E_{\rm F}$ lies in the gap between disorder-broadened Landau levels, there is no percolating network, and currents at the edge of the sample are carried by *edge states* localized along the confining potential region[15]. The percolating network, being infinite in extent in the thermodynamic limit, must connect to the edges. We shall have much more to say about edge states below, but for the moment it is important to apprehend the picture described by Fig. 2.4. Directional edge currents, indicated by the $\odot$ and $\otimes$ symbols in the figure, are responsible for any electrical conduction processes[16].

2.1.5 Scaling of transport data at the IQH transition

The classical picture of the continuum percolation transition is missing something important: quantum tunneling at the saddle points. We know this must be true, but moreover we can

[14]The Hall bar is of course connected to leads for current source and drain, as well as for measuring longitudinal and transverse voltage drops. See Fig. 1.1.

[15]Indeed, this is true whenever $E_{\rm F}$ lies in a *mobility gap*, which is to say whenever $E_{\rm F}$ does not coincide with the continuum percolation threshold. The latter coincidence applies only for energies belonging to a discrete set of values, of measure zero.

[16]At least within linear response theory.

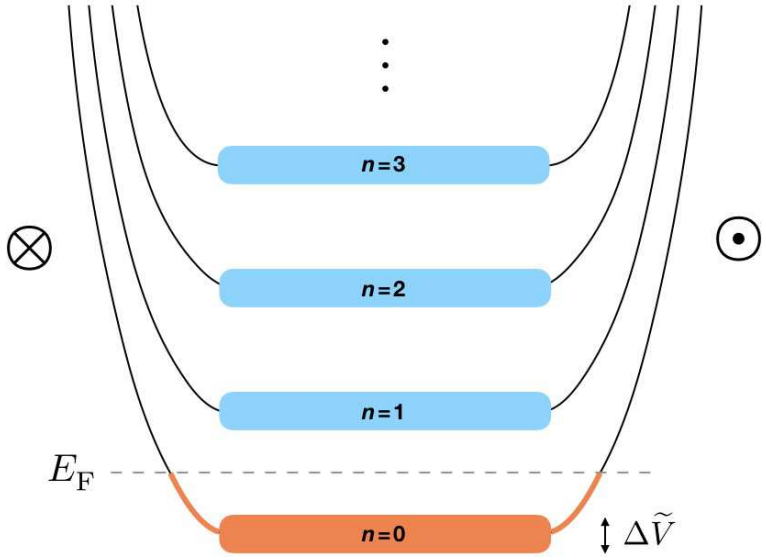


Figure 2.4: Broadened Landau levels and their edge states in a Hall bar. The confining potential forces the energy levels to rise at the edges. When the Fermi level lies between disorder-broadened Landau levels, one or more *edge states* are occupied. The edge states carry the Hall current. The direction of the edge state currents is shown with the $\odot$ and $\otimes$ symbols. Continuum percolation applies when the Fermi energy E_F lies in the vicinity of the center of a disorder-broadened Landau level.

actually *see* it in experiments[17] of the integer quantum Hall (IQH) transition, such as shown in Fig. 2.5. At the lowest temperatures, the Hall conductance/resistance as a function of magnetic field resembles a step function, as B passes through the critical value B_n where $(n+\frac{1}{2})\hbar\omega_c$ passes through the Fermi level. In the vicinity of these critical fields, the correlation length for electrons at the Fermi level diverges as $\xi(B) \propto |B - B_n|^{-\nu}$, where ν is the correlation length exponent. Recall that for classical percolation in $d = 2$ dimensions, $\nu = \frac{4}{3}$ (an exact result). It is then natural to adopt the *scaling hypothesis*

$$\sigma_{yx}(B, L) = \frac{ne^2}{h} + \frac{e^2}{h} F_{\pm}(L/\xi) \tag{2.26}$$

in a system of linear dimension L, where $F_{\pm}(u)$ are scaling functions for $B \gtrless B_n$ with values $F_{\pm}(0) = \frac{1}{2}$ for the transition between consecutive Landau levels[18], $F_{-}(\infty) = 0$ and $F_{+}(\infty) = 1$. The functions $F_{\pm}(u)$ are presumed to interpolate smoothly between their limiting values at $u = 0$ and $u = \infty$. In the thermodynamic limit $L \gg \xi \gg 1$, we have $\sigma_{yx}(B, L \to \infty) = \big(n + \Theta(B - B_n)\big)\, e^2/h$, but with finite L, the step is rounded. Finite temperature T plays a

[17] See, *e.g.*, W. Li *et al.*, *Phys. Rev. Lett.* **102**, 216801 (2009).
[18] This is per spin degree of freedom, or assuming complete spin polarization.

similar role to finite length. With $\beta = 1/k_{\rm B}T$, the scaling *Ansatz* takes the form[19]

$$\sigma_{yx}(B, L, T) = \frac{e^2}{h}\, F(L/\xi, \beta/\xi_\tau) = \frac{e^2}{h}\, \widetilde{F}\big(L^{1/\nu}\,|B - B_n|\,, T^{-1/\nu z}\,|B - B_n|\big) \quad , \tag{2.27}$$

where $\xi_\tau = \xi^z \propto |B - B_n|^{-\nu z}$, where z is the *dynamic critical exponent*. We will see that $z = 1$ for the IQH transition; indeed this can be inferred by comparing transport data as functions of L and of T[20]. The condition $z = 1$ is generally believed to apply in the presence of Coulomb interactions, where the energy scale near criticality is $\hbar\omega \sim e^2/\xi$. For strictly noninteracting and nonrelativistic systems, however, $z = 2$. The reason is that the spectral function $S(\boldsymbol{q}, \omega; E)$ can be expressed using current conservation in terms of the diffusion coefficient $D(\boldsymbol{q}, \omega; E)$ as[21]

$$S(\boldsymbol{q}, \omega; E) = \frac{\hbar\rho(E)}{\pi}\, \frac{q^2 D(\boldsymbol{q}, \omega; E)}{\omega^2 + \big(\boldsymbol{q}^2\, D(\boldsymbol{q}, \omega; E)\big)^2} \quad , \tag{2.28}$$

where $\rho(E)$ is the single particle density of states, with units of $E^{-1}L^{-2}$. If the system is scale invariant at the critical energy $E_{\rm c}$, then $D(\boldsymbol{q}, \omega; E_{\rm c})$ can depend only on the dimensionless combination qL_ω where $L_\omega = \big(\rho(E)\,\hbar\omega\big)^{-1/2}$. Thus if $\rho(E_{\rm c})$ is finite, then $S(\boldsymbol{q}, \omega; E_{\rm c})$ is a function of the combination ω/q^2, which is equivalent to $z = 2$. The takeaway point here is that *interactions must be invoked if the experimentally supported result $z \approx 1$ is to be explained*. We shall return to this point later on below.

Consider now the derivative $d\sigma_{yx}/dB$ as one goes through the IQH transition,

$$\frac{\partial\sigma_{yx}}{\partial B} = \frac{e^2}{h}\, L^{1/\nu}\, \widetilde{F}_u(u, v) + \frac{e^2}{h}\, T^{-1/\nu z}\, \widetilde{F}_v(u, v) \quad , \tag{2.29}$$

where $u \equiv L^{1/\nu}\,|B - B_n|$, $v \equiv T^{-1/\nu z}\,|B - B_n|$, $\widetilde{F}_u = \partial\widetilde{F}/\partial u$, and $\widetilde{F}_v = \partial\widetilde{F}/\partial v$. It is natural to assume that the scaling function $\widetilde{F}(u, v)$ is separately monotonic in each of its arguments, and that the maximum value of $\partial\sigma_{yx}/\partial B$ will occur at $B = B_n$, where $u = v = 0$. Thus,

$$\left(\frac{\partial\sigma_{yx}}{\partial B}\right)_{\rm max} = \frac{e^2}{h}\, \big(c_1\, L^{1/\nu} + c_2\, T^{-1/\nu z}\big) \quad , \tag{2.30}$$

where $c_1 = \widetilde{F}_u(0, 0)$ and $c_2 = \widetilde{F}_v(0, 0)$ are dimensionful constants. In the zero temperature or thermodynamic limits, where we take $T = 0$ or $L = \infty$ at the start, in which case

$$\begin{aligned} \left(\frac{\partial\sigma_{yx}}{\partial B}\right)_{\rm max}(L, 0) &= \frac{e^2}{h}\, d_1\, L^{1/\nu} \\ \left(\frac{\partial\sigma_{yx}}{\partial B}\right)_{\rm max}(\infty, T) &= \frac{e^2}{h}\, d_2\, T^{-1/\nu z} \quad , \end{aligned} \tag{2.31}$$

[19] Here we suppress the $\pm$ indices on the scaling function for notational convenience.

[20] Among the earliest works to report $(d\rho_{xy}/dB)_{\rm max} \sim T^{-\kappa}$ with $\kappa = 1/\nu z = 0.42$ is H. P. Wei *et al.*, *Phys. Rev. Lett.* **61**, 1294 (1988). Samples of different width W were systematically studied by S. Koch *et al.*, *Phys. Rev. Lett.* **67**, 883 (1991), who found that the half-width of the $\rho_{xy}(B, T)$ at the lowest temperatures scaled as $\Delta B \sim W^{-1/\nu}$ with $\nu = 2.34 \pm 0.04$.

[21] See the review by B. Huckestein (RMP, 1995).

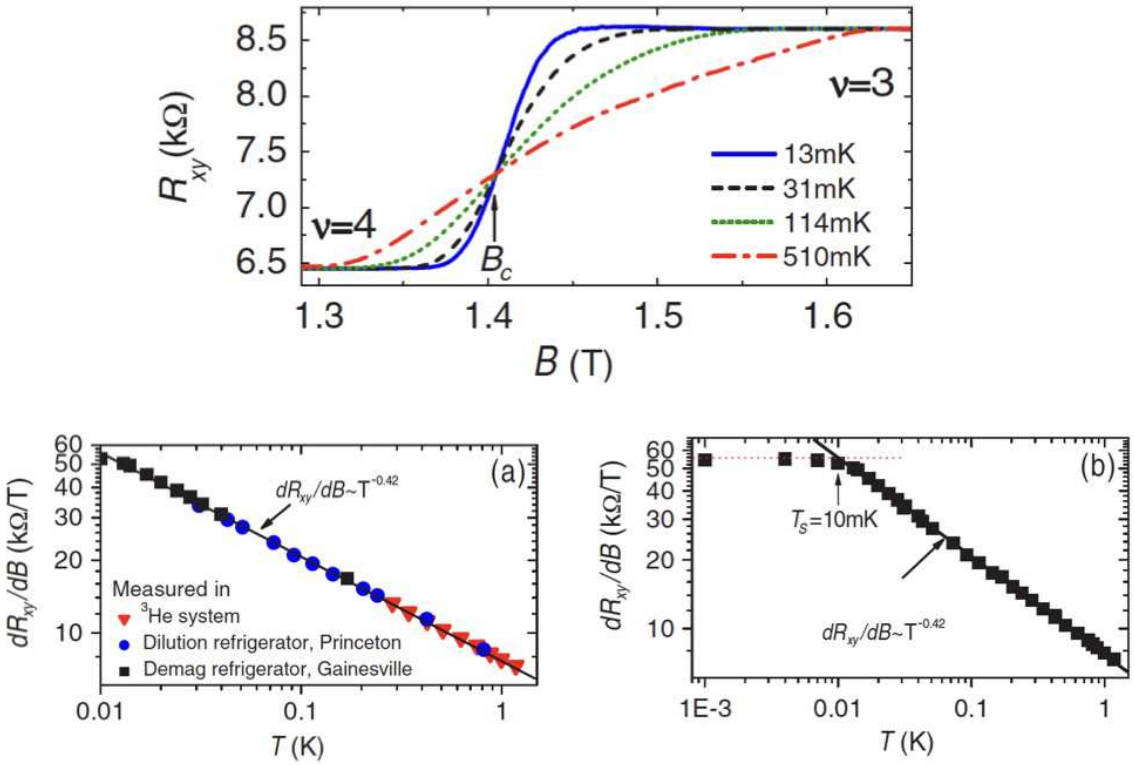


Figure 2.5: Temperature scaling of the $\nu = 3$ to $\nu = 4$ integer quantum Hall transition. Data are from W. Li *et al.*, *Phys. Rev. Lett.* **102**, 216801 (2009). At low temperatures, down to approximately $T = 10\,\mathrm{mK}$, the maximum value of dR_{xy}/dB scales as $T^{0.42} = T^{1/z\nu}$, where $z = 1$ is the dynamical critical exponent and $\nu \approx 2.35$ is the correlation length exponent for the QHE transition.

where $d_1 = \widetilde{F}_u(0,\infty)$ and $d_2 = \widetilde{F}_v(\infty,0)$. Note that while the maximum slope in the $L = \infty$ and $T = 0$ limits is infinite (the derivative of a step function), in each case this infinity is blunted, with the maximum slope being proportional to $L^{1/\nu}$ or $T^{-1/\nu z}$.

In the experiments of Li *et al.*, results of which are shown in Fig. 2.5, the $\nu = 3$ to $\nu = 4$ IQHE transition was observed in transport for temperatures roughly between $T = 1\,\mathrm{mK}$ and $T = 1\,\mathrm{K}$. The sample widths ranged from $W = 100\,\mu\mathrm{m}$ to $W = 500\,\mu\mathrm{m}$. At the critical field $B_\mathrm{c} \approx 1.4\,\mathrm{T}$, the magnetic length is $\ell = 217\,\text{Å}$, so $W/\ell \sim 10^4$. The maximum value of dR_{xy}/dB was found to scale with temperature as $T^{0.42}$ down to $T \approx 10\,\mathrm{mK}$, below which it remained fixed. This latter behavior is associated with finite size effects, *i.e.* the regime $\xi(B) > L$. Over the scaling regime, setting $1/\nu z \simeq 0.42$, one obtains $\nu z \simeq 2.38$. Again, it will turn out that $z = 1$, as can be inferred from size dependence of $(dR_{xy}/dB)_{\max}$, in which case $\nu \simeq 2.38$.

A more general form of the scaling *Ansatz*, for a physical quantity Γ, is

$$\Gamma(B, L, T, \boldsymbol{k}, \omega, \ldots) = \xi^{y_\Gamma}\, F\big(L/\xi, T\xi^{-z}, \boldsymbol{k}\xi, \omega/T, \ldots\big) \tag{2.32}$$

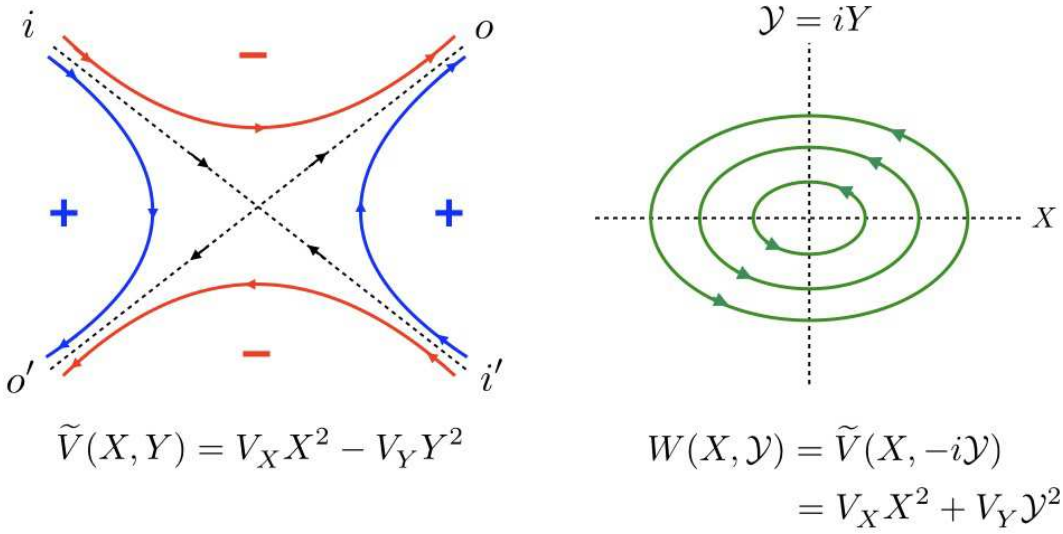


Figure 2.6: Quantum tunneling across saddle points in a random potential $\widetilde{V}(\boldsymbol{r})$. Left: Saddle point in the (X, Y) plane. Right: Upon analytic continuation to imaginary space $\check{Y} = iY$, the saddle becomes a center.

where $\xi \propto |B - B_n|^{-\nu}$ is the correlation length, and where y_Γ is the *scaling dimension* of Γ. The Hall conductivity σ_{xy} has scaling dimension $y = 0$.

A related approach to the temperature scaling is to invoke the notion of an *inelastic scattering length* which diverges as $\ell_{\rm in} \sim T^{-p/2}$ in the limit $T \to 0$. In this picture[22], the scaling variables are ξ/L and $\xi/\ell_{\rm in} \propto \left(T^{-p/2\nu}|B - B_{\rm c}|\right)^{-\nu}$. Thus, in the thermodynamic limit $L \to \infty$, we have $(d\rho_{xy}/dB)_{\rm max} \propto T^{-\kappa}$ with $\kappa = p/2\nu$.

2.1.6 Quantum tunneling across saddle points

Assuming the validity of the scaling *Ansatz*, the transport data are inconsistent with $d = 2$ continuum percolation, for which $\nu = \frac{4}{3}$. What is missing, of course, is quantum mechanics. The problem of tunneling across a saddle point of an electron in a high magnetic field was considered by Fertig and Halperin[23] and by Jain and Kivelson[24]. Consider the situation in the left panel of Fig. 2.6. There are two incoming channels, marked i and i', and two outgoing channels, marked o and o'. The S-matrix acts on incoming flux amplitudes to yield outgoing

[22] See A. M. M. Pruisken, *Phys. Rev. Lett.* **61**, 1297 (1988).
[23] H. A. Fertig and B. I. Halperin, *Phys. Rev. B* **36**, 7969 (1987).
[24] J. K. Jain and S. Kivelson, *Phys. Rev. B* **37**, 4111 (1988).

flux amplitudes, *viz.*

$$\begin{pmatrix} o' \\ o \end{pmatrix} = \overbrace{\begin{pmatrix} r & t' \\ t & r' \end{pmatrix}}^{\mathcal{S}} \begin{pmatrix} i \\ i' \end{pmatrix} \quad , \tag{2.33}$$

where r, r' are reflection amplitudes and t, t' are transmission amplitudes. The reflection and transmission *probabilities* are given by the squares of the corresponding amplitudes,

$$R = |r|^2 \quad , \quad T = |t|^2 \quad , \quad R' = |r'|^2 \quad , \quad T' = |t'|^2 \quad . \tag{2.34}$$

Unitarity of $\mathcal{S}$ guarantees that $R' = R$ and $T' = T$, as well as $R + T = 1$. Note that $\mathcal{S} \in \mathsf{U}(2)$, which has dimension four. This allows us to write $t' = t^* \exp(-i\delta)$ and $r' = -r^* \exp(-i\delta)$ with $r = \sin(\theta)\exp(i\psi)$ and $t = \cos(\theta)\exp(i\omega)$. The four parameters are then $(\theta, \psi, \omega, \delta)$.

It is important to recognize that the complex scalars $\{i, i', o, o'\}$ are *flux amplitudes* and not wavefunction amplitudes. Unitarity of $\mathcal{S}$ means that

$$|o'|^2 + |o|^2 = |i|^2 + |i'|^2 \quad , \tag{2.35}$$

and is a statement about *current conservation*. A simple illustration of the difference is afforded by consideration of the one-dimensional step potential $V(x) = V_0\,\Theta(x)$. We write

$$\begin{aligned} x < 0 \ : \quad & \psi(x) = I\,e^{ikx} + O'\,e^{-ikx} \\ x > 0 \ : \quad & \psi(x) = O\,e^{ik'x} + I'\,e^{-ik'x} \quad , \end{aligned} \tag{2.36}$$

where the energy

$$E = \frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 k'^2}{2m} + V_0 \tag{2.37}$$

is conserved in by the scattering process, and is assumed to be positive. The requirements that $\psi(x)$ and $\psi'(x)$ be continuous at $k = 0$ provide two conditions on the four wavefunction amplitudes:

$$\begin{aligned} I + O' &= O + I' \\ k\,(I - O') &= k'\,(O - I') \quad . \end{aligned} \tag{2.38}$$

The flux amplitudes $\{i, i', o, o\}$ are related to the wavefunction amplitudes $\{I, I', O, O'\}$ by a multiplicative factor of the square root of the velocity, where $v = \hbar k/m$ and $v' = \hbar k'/m$:

$$\begin{pmatrix} i \\ o' \end{pmatrix} = \sqrt{v} \begin{pmatrix} I \\ O' \end{pmatrix} \qquad , \qquad \begin{pmatrix} o \\ i' \end{pmatrix} = \sqrt{v'} \begin{pmatrix} O \\ I' \end{pmatrix} \quad . \tag{2.39}$$

One may now derive the $\mathcal{S}$-matrix,

$$\mathcal{S} = \begin{pmatrix} r & t' \\ t & r' \end{pmatrix} = \frac{1}{1+\eta} \begin{pmatrix} 1-\eta & 2\sqrt{\eta} \\ 2\sqrt{\eta} & \eta - 1 \end{pmatrix} \quad , \tag{2.40}$$

where

$$\eta = \frac{v'}{v} = \frac{k'}{k} = \sqrt{1 - \frac{V_0}{E}} \quad . \tag{2.41}$$

One can check that $\mathcal{S}^\dagger \mathcal{S} = 1$. However note that the matrix $\widetilde{\mathcal{S}}$ which acts on the wavefunction amplitudes, with

$$\begin{pmatrix} O' \\ O \end{pmatrix} = \widetilde{\mathcal{S}} \begin{pmatrix} I \\ I' \end{pmatrix} \tag{2.42}$$

is related to $\mathcal{S}$ by

$$\widetilde{\mathcal{S}} = \begin{pmatrix} 1/\sqrt{v'} & 0 \\ 0 & 1/\sqrt{v} \end{pmatrix} \mathcal{S} \begin{pmatrix} \sqrt{v} & 0 \\ 0 & \sqrt{v'} \end{pmatrix} \tag{2.43}$$

and is in general not unitary. Note also that when $v = v'$ we have $\widetilde{\mathcal{S}} = \mathcal{S}$.

Saddle point transmission probability

For an electron in the potential

$$V(x, y) = V_x \, x^2 - V_y \, y^2 \quad , \tag{2.44}$$

Fertig and Halperin (1987) obtained the transmission probability[25]

$$T(\epsilon) = \frac{1}{1 + \exp(\pi\epsilon)} \quad , \tag{2.45}$$

where $\epsilon = \big(E - (n + \frac{1}{2})\hbar\omega_{\rm c}\big)/\Gamma$ and $\Gamma \approx \ell^2 (V_x V_y)^{1/2}$, assuming $\ell^2 \, V_{x,y} \ll \hbar\omega_{\rm c}$. Note that for $|\epsilon| \gg 1$ one has

$$T(\epsilon) \simeq \begin{cases} \exp\big(-\pi|\epsilon|\big) & \text{if } \epsilon > 0 \\ 1 - \exp\big(-\pi|\epsilon|\big) & \text{if } \epsilon < 0 \end{cases} \quad . \tag{2.46}$$

Similarly,

$$R(\epsilon) = 1 - T(\epsilon) \simeq \begin{cases} 1 - \exp\big(-\pi|\epsilon|\big) & \text{if } \epsilon > 0 \\ \exp\big(-\pi|\epsilon|\big) & \text{if } \epsilon < 0 \end{cases} \quad . \tag{2.47}$$

This conforms to the situation in Fig. 2.6: If $\epsilon \gg 1$, the transmission is almost purely from i to o' and from i' to o, corresponding to $R \approx 1$ and $T \approx 0$. If on the other hand $\epsilon \ll -1$, then the transmission is almost purely from i to o and from i' to o', hence $R \approx 0$ and $T \approx 1$.

[25] For us, the parameter ϵ is the negative of that in Fertig and Halperin.

Instantons and tunnel splittings

Following Jain and Kivelson (1988), we can also apply the coherent state path integral to this problem. The partition function at inverse temperature $\beta = 1/k_{\rm B}T$ is given by

$$Z(\beta) = \mathsf{Tr}\, e^{-\beta H} = \int \frac{d^2R}{2\pi\ell^2}\, \langle\, \boldsymbol{R}\,|\, e^{-\beta H}\,|\,\boldsymbol{R}\,\rangle = \int\limits_{\boldsymbol{R}(0)=\boldsymbol{R}(\hbar\beta)} D\big[R(\tau),\bar{R}(\tau)\big]\; e^{-\mathcal{S}_{\rm E}/\hbar} \quad . \tag{2.48}$$

The partition function is the Laplace transform of the density of states, and as such incorporates all information about the energy spectrum, including tunnel splittings. We measure energies with respect to $(n+\frac{1}{2})\hbar\omega_{\rm c}$, in which case

$$\mathcal{S}_{\rm E}\big[X(\tau),Y(\tau)\big] = \int\limits_0^{\hbar\beta} d\tau\, \left[\frac{i\hbar}{2\ell^2}\big(Y\dot{X} - X\dot{Y}\big) + \widetilde{V}(X,Y)\right] + \Delta\mathcal{S}_{\rm E} \quad , \tag{2.49}$$

where $\widetilde{V}(X,Y) \equiv \widetilde{V}(R|\bar{R}) = \langle\, \boldsymbol{R}\,|\,V\,|\,\boldsymbol{R}\,\rangle$ with $R = X + iY$ and $\bar{R} = X - iY$, and where $\Delta\mathcal{S}_{\rm E}$ is the boundary discontinuity term. The equations of motion obtained by extremizing $\mathcal{S}_{\rm E}$ force us to analytically continue to imaginary *space*. In terms of the complexified guiding center coordinates R and $\bar{R}$, this entails $\bar{R} \neq R^*$ along the instanton path. Writing $\mathcal{Y} \equiv iY$[26] and defining $W(X,\mathcal{Y}) \equiv \widetilde{V}(X,-i\mathcal{Y})$, we obtain $\mathcal{S}_{\rm E} = \int_0^{\hbar\beta} d\tau\; L(X,\mathcal{Y},\dot{X},\dot{\mathcal{Y}}) + \Delta\mathcal{S}_{\rm E}$, where the Lagrangian is

$$L = \frac{\hbar}{2\ell^2}\big(\mathcal{Y}\dot{X} - X\dot{\mathcal{Y}}\big) + W(X,\mathcal{Y}) \quad . \tag{2.50}$$

The equations of motion are then

$$\dot{X} = -\frac{\ell^2}{\hbar}\,\frac{\partial W}{\partial \mathcal{Y}} \qquad , \qquad \dot{\mathcal{Y}} = \frac{\ell^2}{\hbar}\,\frac{\partial W}{\partial X} \quad . \tag{2.51}$$

Note that W is then conserved along the trajectory, since

$$\frac{d}{dt}\,W\big(X(t),\mathcal{Y}(t)\big) = \frac{\partial W}{\partial X}\,\dot{X} + \frac{\partial W}{\partial \mathcal{Y}}\,\dot{\mathcal{Y}} = 0 \quad . \tag{2.52}$$

Assuming the boundary discontinuity term vanishes, the Euclidean action is then

$$\mathcal{S}_{\rm E} = \beta E + \frac{A}{\ell^2} \tag{2.53}$$

where E is the conserved value of the potential along the instanton trajectory and A is the area enclosed by the trajectory. For the saddle point potential $\widetilde{V}(X,Y) = V_X X^2 - V_Y Y^2$, we have $W(X,\mathcal{Y}) = V_X X^2 + V_Y \mathcal{Y}^2$, and setting $W = E$ we obtain the ellipse

$$\frac{X^2}{a^2} + \frac{\mathcal{Y}^2}{b^2} = 1 \quad , \tag{2.54}$$

[26] Note we have repurposed the symbol $\mathcal{Y}$ here.

with $a = (E/V_X)^{1/2}$ and $b = (E/V_Y)^{1/2}$. The full area of the ellipse is $A = \pi ab = \pi E/(V_X V_Y)^{1/2}$, hence $A/\ell^2 = \pi\epsilon$, where $\epsilon = E/\Gamma$. Recall that here we measure E relative to the center of the Landau level at $E_n = (n + \frac{1}{2})\hbar\omega_c$.

For weak tunneling, the amplitude t is proportional to the single instanton contribution after subtracting off the βE term, and is given by $t = \exp(-A/2\ell^2)$ since only half of the elliptical trajectory of $(X, \mathcal{Y})$ is traversed in crossing the saddle. The transmission coefficient is then given by $T = |t|^2 = \exp(-A/\ell^2) = \exp(-\pi\epsilon)$, exactly as in Fertig and Halperin. For smaller values of ϵ, multiple instanton paths must be summed over in order to get the Fertig-Halperin result $T(\epsilon) = \left[1 + \exp(\pi\epsilon)\right]^{-1}$.

Mil'nikov-Sokolov argument

G. Mil'nikov and I. Sokolov (1988) published a seductive argument[27] purporting to establish that the correlation length exponent for quantum continuum percolation should be given by $\nu_{QU} = \nu_{CL} + 1 = \frac{7}{3}$. Consider a distance $r \gg \xi_{CL}(E)$, over which one expects to encounter $\sim r/\xi_{CL}(E)$ such saddles. The probability of transmission across this entire distance is then

$$T_{QU}(E) \sim \left(T_{saddle}(E)\right)^{r/\xi_{CL}(E)} = e^{-\pi r|E|/\Gamma\xi_{CL}(E)} \equiv e^{-r/\xi_{QU}(E)} \quad . \tag{2.55}$$

Thus, we expect

$$\xi_{QU}(E) = \langle\Gamma\rangle \frac{\xi_{CL}(E)}{\pi|E|} \propto |E|^{-7/3} \quad , \tag{2.56}$$

where $\langle\Gamma\rangle$ is an average over the Γ parameter over many saddles. As we shall see, this result is quite close to the experimentally determined value of $\nu = 2.35$ (see Fig. 2.5). It is also close to the earliest numerical simulation values of ν_{QU} for the IQH transition. Alas, the Mil'nikov-Sokolov argument is bogus – at least insofar as it purports to describe the critical properties of the disordered noninteracting 2DEG in the LLL – because it doesn't properly account for the connectivity of the saddle point network, treating it instead as a chain with no closed loops. It also doesn't account for quantum interference effects associated with backscattering from saddle points. Below we shall see how a more sophisticated treatment, the Chalker-Coddington *network model of quantum percolation*, can properly model the critical behavior of noninteracting electrons in a magnetic field and a random potential.

Away from the quantum critical point near the center of the Landau level – but not too far away – the Mil'nikov-Sokolov picture should be applicable[28]. In this regime, the quantum tunneling probability $e^{-\pi|E|/\Gamma} \ll 1$ is weak, and we can consider only transmission across each saddle, with no closed loops associated with reflection paths. Recall $\Gamma = \ell^2 (V_x V_y)^{1/2} \propto W\ell^2/d^2$, where W is the RMS potential fluctuation and d the correlation length of the potential, assuming there

[27] G. V. Mil'nikov and I. M. Sokolov, *Pis'ma Zh. Eksp. Teor. Fiz.* **48**, 494 (1988) [*JETP Lett.* **48**, 536 (1988)].

[28] See M. M. Fogler, A. Y. Dobin, and B. I. Shklovskii, *Phys. Rev. B* **57**, 4614 (1998). I thank my colleague Misha Fogler for explaining this to me.

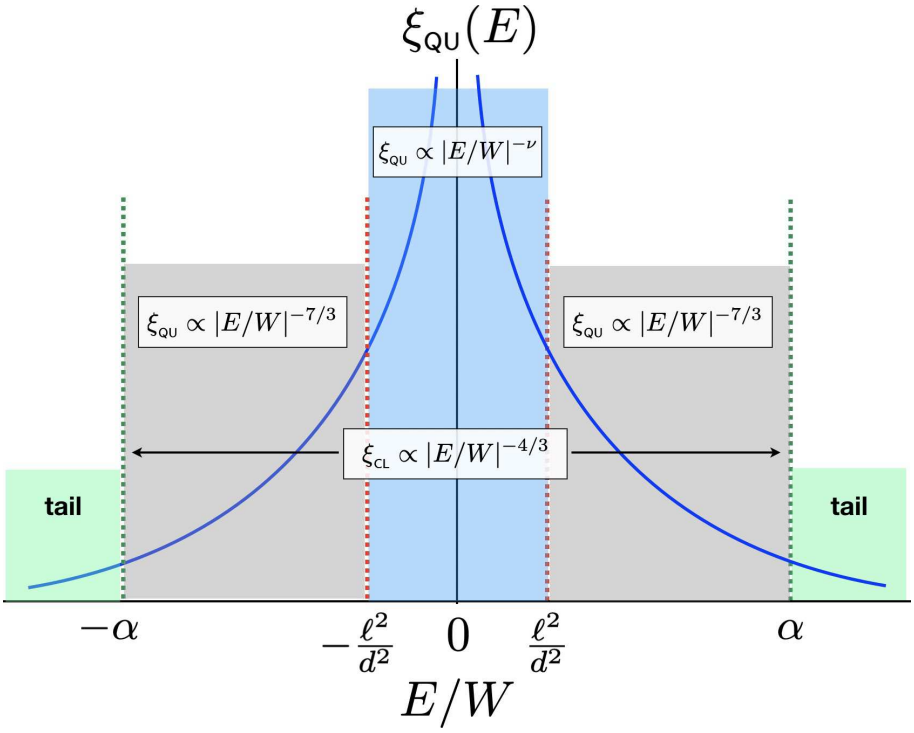


Figure 2.7: Regimes of behavior for $\xi_{CL}(E)$ and $\xi_{QU}(E)$ in a smooth random potential $V(\boldsymbol{r})$ whose RMS fluctuations are W and whose correlation length is d. Blue: Critical regime, in which the exponent ν for quantum percolation is obtained from network model simulations. Gray: Mil'nikov-Sokolov regime, in which quantum tunneling across saddle points is weak but the classical percolation is still in the critical regime. Green: Tail regime, in which the physics is dominated by local fluctuations of $V(\boldsymbol{r})$.

is a single length scale associated with $V(\boldsymbol{r})$, which is the case if $V(\boldsymbol{r})$ is chosen according to the distribution functional

$$P\big[V(\boldsymbol{r})\big] = P[0]\exp\left\{-\frac{1}{2W^2d^2}\int d^2r\,\Big[V^2 + d^2(\boldsymbol{\nabla}V)^2\Big]\right\} \quad . \tag{2.57}$$

Classical continuum percolation then says $\xi_{CL}(E) = Cd\,|E/W|^{-4/3}$, where C is a dimensionless constant. This form is valid provided $|E/W|$ is sufficiently small, which is to say within the critical regime, which is to say $|E| < \alpha W$, where $\alpha = \mathcal{O}(1)$ is a dimensionless constant. The condition that reflections may be neglected is tantamount to $|E| \gtrsim \Gamma$, and thus the range over which we expect the MS argument to be valid is

$$(\ell/d)^2 \lesssim |E|/W \lesssim \alpha \quad . \tag{2.58}$$

For potentials which are very smooth on the scale of ℓ, this criterion is easily satisfied. In this regime, we therefore expect $\xi_{QU}(E) \propto |E|^{-7/3}$. As we shall see, this exponent is quite close to

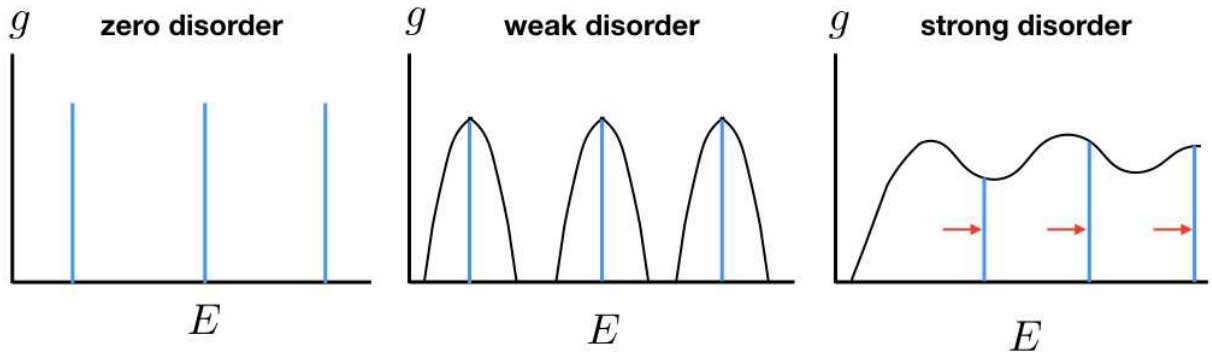


Figure 2.8: Floating up of extended states in the presence of disorder.

what is observed in scaling analyses of experiments, but is inconsistent with the most recent best results for the correlation length exponent for quantum percolation networks.

2.1.7 Landau level mixing and "floating" of extended states

At high fields, each Landau level carries one unit of Hall conductivity e^2/h, *i.e.* the $n^{\rm th}$ Landau level carries Chern number $\Delta C_n = 1$ (or $\Delta C_n = -1$ as in our case with $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$). If we ignore Landau level mixing, this state of affairs persists to weak fields as well, since each LL remains independent. However, in the $B = 0$ limit, in a two-dimensional disordered system with no interactions, all electronic states are known to be localized and the system is an integer. As $B \to 0$, the cyclotron gap between LLs becomes smaller, tending to zero, and eventually we know LL mixing must apply. What happens to all the extended states lying at the LL centers[29]? An early view suggested that extended states must *float up* in energy as $B \to 0$ in the presence of disorder. A cartoon of this notional state of affairs is depicted in Fig. 2.9(a). According to this picture, one should expect reentrant behavior in $\sigma_{xy}(n, B)$ as a function of B at fixed density n. Note how in this picture, there are direct transitions only between states with $\Delta C = \pm 1$, and none for $|\Delta C| > 1$.

What actually occurs, both in numerical simulations as well as in experiments, appears to be more complicated and as yet not fully understood. The fate of extended states at weak disorder was investigated numerically by Sheng, Weng, and Wen (SWW), using a tight binding model for spin-polarized electrons.

$$H = -\sum_{\langle \boldsymbol{r}\boldsymbol{r}'\rangle} \left(e^{iA_{\boldsymbol{r}\boldsymbol{r}'}}\, c^\dagger_{\boldsymbol{r}}\, c_{\boldsymbol{r}'} + e^{-iA_{\boldsymbol{r}\boldsymbol{r}'}}\, c^\dagger_{\boldsymbol{r}'}\, c_{\boldsymbol{r}} \right) + \sum_{\boldsymbol{r}} W_{\boldsymbol{r}}\, c^\dagger_{\boldsymbol{r}}\, c_{\boldsymbol{r}} \quad , \tag{2.59}$$

where the flux per plaquette is taken to be $\phi = 2\pi p/q$ local disorder potential at lattice site $\boldsymbol{r}$ is

[29] Due to asymmetry in the distribution of $\widetilde{V}(\boldsymbol{r})$ and LL mixing, the extended states are not obliged to lie exactly at $E_n = (n + \frac{1}{2})\hbar\omega_{\rm c}$.

given by

$$W_{\boldsymbol{r}} = \frac{W}{\pi} \sum_{\boldsymbol{r}'} f_{\boldsymbol{r}'}\, e^{-|\boldsymbol{r}-\boldsymbol{r}'|^2/\lambda^2} \quad , \tag{2.60}$$

where on each site $f_{\boldsymbol{r}}$ is uniformly distributed on the interval $[-1, 1]$. Thus W and λ are the strength and correlation length of the random disorder. SWW's results are summarized in Fig. 2.9(b). The maximum lattice size was 32×64. The transition to the insulating state occurs along the dark black curve, thus direct transitions were observed from each of $C = 1, 2, 3, 4$ to $C = 0$.

Experiments by Kravchenko *et al.* in Si MOSFETS, shown in Fig. 2.9(c,d), show an apparent violation of the $|\Delta C| = 1$ rule for the IQH transitions inferred from the cartoon picture. Rather, there are a sequence of direct transitions from IQH states with $C = 1$, $C = 2$, $C = 4$, and $C = 6$ to a state they identify as an "insulator". Note how the $C = 3$ and $C = 5$ states get crowded out as disorder increases but before one reaches the $C = 0$ insulator. Thus, there are direct transitions observed between $C = 2$ and $C = 4$ and between $C = 4$ and $C = 6$. These $|\Delta C| = 2$ transitions may be associated with spin-orbit effects in the presence of interactions, although this is more likely to pertain in GaAs heterojunctions where the spin-orbit interaction is stronger than in Si due to larger nuclear Z.

However, there is a rather severe problem with Kravchenko *et al.*'s interpretation of their "insulator"[30]. Consider their results for the electron scattering rate τ^{-1}, which they extract from the $B = 0$ expression for the mobility $\mu = e\tau/m^*$. Using $m^* = (m_{\rm t}^2\, m_{\rm l})^{1/3} = 0.22\, m_{\rm c}$ for the conduction electrons in Si[31], if we divide the Drude formula for ρ_{xx} by the quantum of resistance h/e^2, we obtain

$$\frac{e^2}{h}\,\rho_{xx} = \frac{m^*}{2\pi n\hbar\tau} = \overbrace{\frac{0.22\, m_{\rm e} \times 10^{14}\,{\rm s}^{-1}}{2\pi \times 1.055 \times 10^{-27}\,{\rm erg}\cdot{\rm s} \times 10^{11}\,{\rm cm}^{-2}}}^{=\,3.32} \times \frac{\tau^{-1}\big[10^{14}\,{\rm s}^{-1}\big]}{n\big[10^{11}\,{\rm cm}^{-2}\big]} \quad . \tag{2.61}$$

Thus, for $n \approx 10^{11}\,{\rm cm}^{-1}$ and $\tau^{-1} \lesssim 10^{14}\,{\rm s}^{-1}$, $\rho_{xx} \sim h/e^2$, which can hardly be identified with an insulator! This state is best identified as a correlated metal.

2.2 Integer Quantum Hall Transition

2.2.1 Introduction

Recall the eigenfunctions in the Landau strip basis discussed in §1.3.5. The geometry is cylindrical, with $x \in \mathbb{R}$ and $y \in [0, L_y]$ with periodic boundary conditions in the y-direction. Choosing

[30] I am grateful to Steven Kivelson for explaining this to me.

[31] $m_{\rm t}^* = 0.082\, m_{\rm e}$ and $m_{\rm l}^* = 1.64\, m_{\rm e}$ are the transverse and longitudinal effective masses for Si conduction electrons.

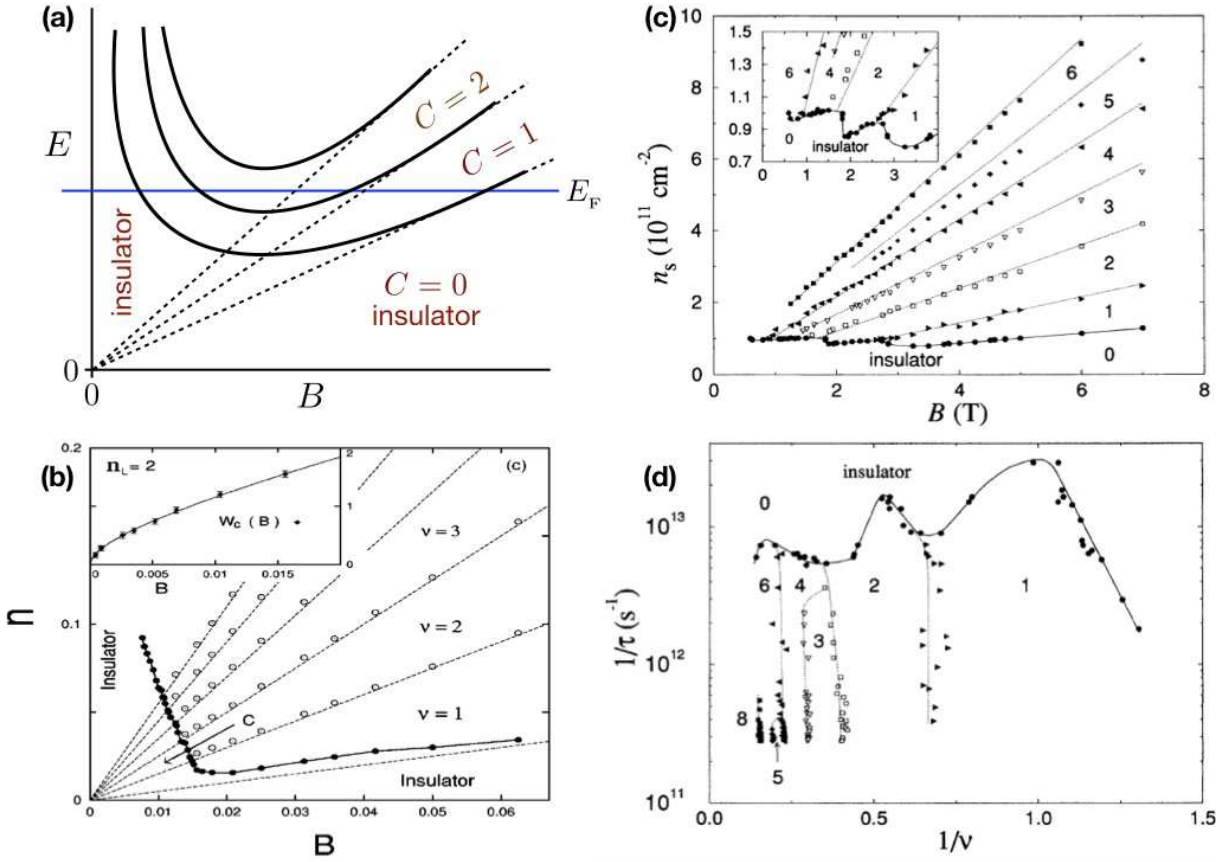


Figure 2.9: Extended and localized states in the quantum Hall effect (a) Early cartoon picture, which suggests reentrant behavior as B is varied at fixed density. Dashed lines correspond to $E = E_n = (n + \frac{1}{2})\hbar\omega_c$. (b) Numerical data of D. N. Sheng, Z. Y. Weng, and X. G. Wen, *Phys. Rev. B* **64**, 165317 (2001). (c) and (d) Experimental data in Si MOSFETS, from S. V. Kravchenko *et al.*, *Phys. Rev. Lett.* **75**, 910 (1995). The scattering rate $1/\tau$ extracted from the data serves as a measure of the strength of the disorder potential $\widetilde{V}(\boldsymbol{r})$. The state labeled as "insulator" in panels (c) and (d) is in fact a correlated metal.

the gauge $\boldsymbol{A} = A_y\,\hat{\boldsymbol{y}}$ with $A_y = -Bx$, we obtained the eigenfunctions

$$\psi_{n,j}(x,y) = L_y^{-1/2}\, e^{ik_y\, y}\, \phi_n(x - \ell^2 k_y) \quad , \tag{2.62}$$

where $E_{n,j} = (n+\frac{1}{2})\hbar\omega_c$ and where $\phi_n(x) = (2^n n!)^{-1/2}(\pi\ell^2)^{-1/4}\mathsf{H}_n(x/\ell)\,\exp(-x^2/2\ell^2)$. Note that $k_y = 2\pi j/L_y$ is quantized according to the PBC $\exp(ik_y L_y) = 1$.

Now consider the gauge $A_y = -Bx + \alpha\phi_0/L_y$ where $\phi_0 = hc/e$ is the Dirac flux quantum and where $\alpha \in \mathbb{R}$ is a dimensionless free parameter. We still have $\boldsymbol{\nabla}\times\boldsymbol{A} = -B\hat{\boldsymbol{z}}$, and it is easy to check that this incorporates the gauge transformation

$$H_0(\alpha) = e^{-2\pi i\alpha y/L_y}\, H_0(0)\, e^{+2\pi i\alpha y/L_y} \quad . \tag{2.63}$$

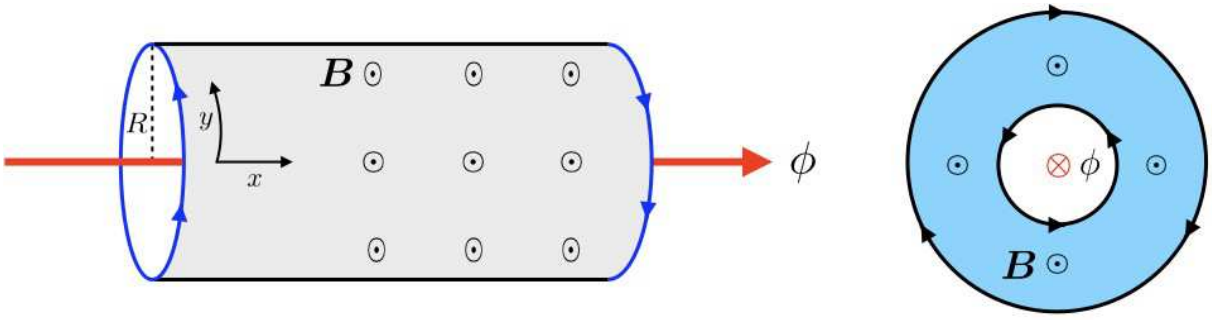


Figure 2.10: Cylindrical and Corbino ring geometries. With each additional fluxoid $\Delta\phi = hc/e$ threaded, an integer number of electrons are transferred from one edge to the other.

One might think that this permits us to write $\psi_{n,j}(x,y;\alpha) = \exp(-2\pi i\alpha y/L_y)\,\psi_{n,j}(x,y;0)$, however such a wavefunction does not satisfy the PBCs except in cases where $\alpha \in \mathbb{Z}$. The resolution is to first shift the k_y quantization such that $\exp(ik_yL_y)\,\exp(-2\pi i\alpha) = 1$, *i.e.* $k_y = 2\pi(j+\alpha)/L_y$. Then

$$\psi_{n,j}(x,y;\alpha) = L_y^{-1/2}\,\exp(2\pi ijy/L_y)\,\phi_n\left(x - \frac{2\pi(j+\alpha)\ell^2}{L_y}\right) \quad , \tag{2.64}$$

Note that

$$\psi_{n,j}(x,y;\alpha+1) = e^{-2\pi iy/L_y}\,\phi_{n,j+1}(x,y;\alpha) \quad , \tag{2.65}$$

which is an allowed transformation, *i.e.* one which preserves the boundary conditions. What has happened (see Fig. 2.10 is that we have threaded our cylinder with α Dirac flux quanta. As we adiabatically increase the flux parameter α by $\Delta\alpha = 1$, the $j^{\rm th}$ Landau strip eigenfunction evolves into the $(j+1)^{\rm th}$ eigenfunction, up to the PBC-preserving gauge factor, $\exp(-2\pi iy/L_y)$.

Now imagine that there is also potential $V_{\rm conf}(x)$ which confines the system in the x-direction. So long as $|\ell\,\nabla V_{\rm conf}| \ll \hbar\omega_{\rm c}$, the eigenfunctions will remain localized along the Landau strips, and the energy eigenvalues will be given by $E_{n,j} = V_{\rm conf}(x_{j,\alpha})$, where $x_{j,\alpha} = 2\pi(j+\alpha)\ell^2/L_y$. If the chemical potentials on the two edges differ by $\Delta\mu = -eV$, where V is the voltage drop, then the current may be computed as

$$I = -c\,\frac{\partial U}{\partial\phi} \approx -c\,\frac{\Delta U}{\phi_0} = c\,\frac{neV}{hc/e} = \frac{ne^2}{h}\,V \quad , \tag{2.66}$$

where U is the total energy, and where n is the number of edge state channels on each side lying below the Fermi level (see Fig. 2.4). This is Laughlin's argument[32] for the quantization of Hall conductance: $\sigma_{yx} = I/V = ne^2/h$. Furthermore, even if the potential $V(\boldsymbol{r})$ is nonzero in the bulk, provided the edge states are localized along the walls of the confining potential, adiabatic increase of the dimensionless flux parameter α by $\Delta\alpha = 1$ still must result in the shifting of edge states by $\Delta j = 1$ for each fully occupied Landau level, and must therefore result in an integer contribution to σ_{yx} in units of e^2/h. The adiabatic change of the threaded

[32] R. B. Laughlin, *Phys. Rev. Lett.* **23**, 5632 (1981).

flux $\Delta\phi = \phi_0$ in the cylindrical or Corbino geometries acts as a pump, transferring n electrons from one side of the sample to the other. Laughlin's argument was subsequently sharpened by Halperin[33].

2.2.2 Replica field theory of the IQH transition

A replica field theory of the transition was proposed in 1983 by Khmel'nitskii[34] and by Libby, Levine, and Pruisken[35], based on the Lagrangian,

$$\mathcal{L}(Q) = \tfrac{1}{4}\,\tilde\sigma^0_{xx}\,\mathsf{Tr}\,(\partial_\mu Q)^2 + \tfrac{1}{8}\,\tilde\sigma^0_{xy}\,\mathsf{Tr}\,\big(\epsilon^{\mu\nu} Q\,\partial_\mu Q\,\partial_\nu Q\big) \quad , \tag{2.67}$$

where $\tilde\sigma^0_{\mu\nu} = h\sigma^0_{\mu\nu}/e^2$ is the dimensionless bare conductivity tensor (*i.e.* at some microscopic length scale), and where $Q(x,y)$ lives in the symmetric coset space $\mathsf{U}(2n)/\mathsf{U}(n)\times\mathsf{U}(n)$ in the $n\to 0$ replica limit. The second term is a topological invariant:

$$\int d^2x\;\mathsf{Tr}\,\big(\epsilon^{\mu\nu} Q\,\partial_\mu Q\,\partial_\nu Q\big) = 16\pi i q \quad , \tag{2.68}$$

where $q\in\mathbb{Z}$; this is true for all n. This field theory is difficult to analyze due to the topological term, but a dilute instanton gas expansion generates the renormalization group flow[36]

$$\begin{aligned}\frac{\partial\tilde\sigma_{xx}}{\partial\ln L} &= -\frac{1}{2\pi^2\tilde\sigma_{xx}} - c\,\tilde\sigma_{xx}\,e^{-2\pi\tilde\sigma_{xx}}\cos(2\pi\tilde\sigma_{xy}) \\ \frac{\partial\tilde\sigma_{xy}}{\partial\ln L} &= -c\,\tilde\sigma_{xx}\,e^{-2\pi\tilde\sigma_{xx}}\sin(2\pi\tilde\sigma_{xy}) \quad .\end{aligned} \tag{2.69}$$

These results are purported to hold at weak coupling $g = 2/\tilde\sigma^0_{xx} \ll 1$. Note that the exponential terms are *nonperturbative* and proportional to $\exp(-4\pi/g)$. What happens for strong coupling as $g\to 0$? We can only guess, consistent with symmetry and physical intuition, and a sketch is shown in Fig. 2.11. The flow in Eqns. 2.69 is valid for large σ^0_{xx}. We know that $\sigma_{xy} = pe^2/h$ must be a stable RG fixed point for $p\in\mathbb{Z}$. This naturally leads to a conjectured unstable fixed point, depicted in red in the figure, and the associated flow lines. We emphasize that the Lagrangian density $\mathcal{L}(Q)$ corresponds to a *two-dimensional* field theory for *noninteracting* electrons.

2.2.3 Chalker-Coddington network model

Despite the elegance of Laughlin's argument and the proposed nonlinear sigma model field-theoretic formulation of the IQH transitions, detailed investigations of the transition, *e.g.* in

[33] B. I. Halperin, *Phys. Rev. B* **25**, 2185 (1982).
[34] D. E. Khmel'nitskii, *Pis'ma Zh. Eksp. Teor. Fiz.* **38**, 454 (1983) [*JETP Lett.* **38**, 552 (1984)].
[35] H. Levine, S. B. Libby, and A. M. M. Pruisken, *Phys. Rev. Lett.* **51**, 1915 (1983).
[36] A. M. M. Pruisken, in *The Quantum Hall Effect*, R. Prange and S. M. Girvin, eds. (Springer, 1987).

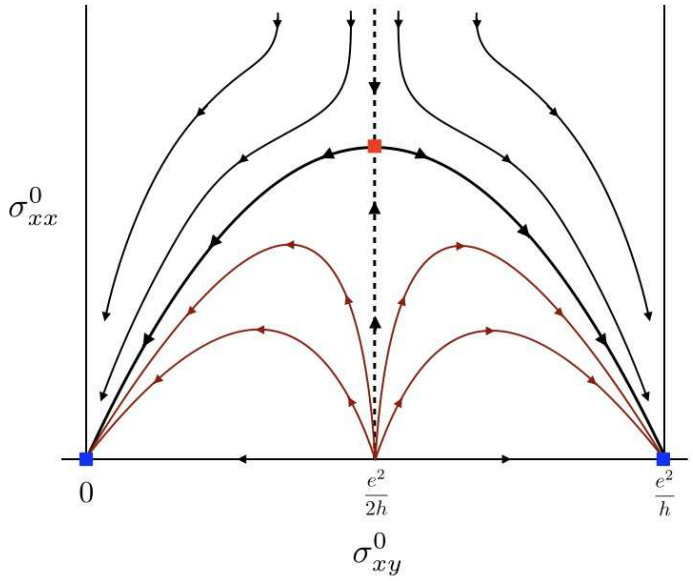


Figure 2.11: Khmel'nitskii-Pruisken RG flow for the nonlinear sigma model field theory of the integer quantum Hall effect. The flow is periodic in σ^0_{xy} with period h/e^2. The conjectured flow at strong coupling is shown in brown. The blue squares represent stable RG fixed points, and the red square an unstable RG fixed point.

terms of critical behavior, is largely problem for numerical simulators. One can, for example, model the an electron in a random potential in a magnetic field with a tight binding model. Such approaches are not optimally suited to gleaning the essential critical behavior because they retain a lot of inessential details, such as regions with high peaks and regions with low valleys, which slow down the computation. A much more efficient model was devised by Chalker and Coddington in 1988[37], who realized that the most important physics is that associated with tunneling across saddle points.

We begin with the single saddle point of Figs. 2.6. The incoming flux amplitudes i and i' scatter into the outgoing amplitudes o and o' according to the linear relation of Eqn. 2.33. As shown in Fig. 2.13, this linear relation between incoming and outgoing amplitudes may be recast, reading the scattering diagram from left to right, as a relation between "left" and "right" amplitudes. Specifically,

$$\begin{pmatrix} o' \\ o \end{pmatrix} = \overbrace{\begin{pmatrix} r & t^* e^{-i\delta} \\ t & -r^* e^{-i\delta} \end{pmatrix}}^{\mathcal{S}} \begin{pmatrix} i \\ i' \end{pmatrix} \qquad \Longrightarrow \qquad \begin{pmatrix} o \\ i' \end{pmatrix} = \overbrace{\begin{pmatrix} 1/t^* & -r^*/t^* \\ -r\, e^{i\delta}/t^* & e^{i\delta}/t^* \end{pmatrix}}^{\mathcal{M}} \begin{pmatrix} i \\ o' \end{pmatrix} . \tag{2.70}$$

The scattering matrix is unitary, satisfying $\mathcal{S}^\dagger \mathcal{S} = \mathbf{1}$. The transfer matrix $\mathcal{M}$ is *pseudo-unitary*, satisfying $\mathcal{M}^\dagger Z \mathcal{M} = Z$, with $Z = \mathrm{diag}(1, -1)$ is the Pauli σ^z matrix. This guarantees $|i|^2 - |o'|^2 =$

[37] J. T. Chalker and P. D. Coddington, *J. Phys. C: Solid State Phys.* **21**, 3665 (1988).

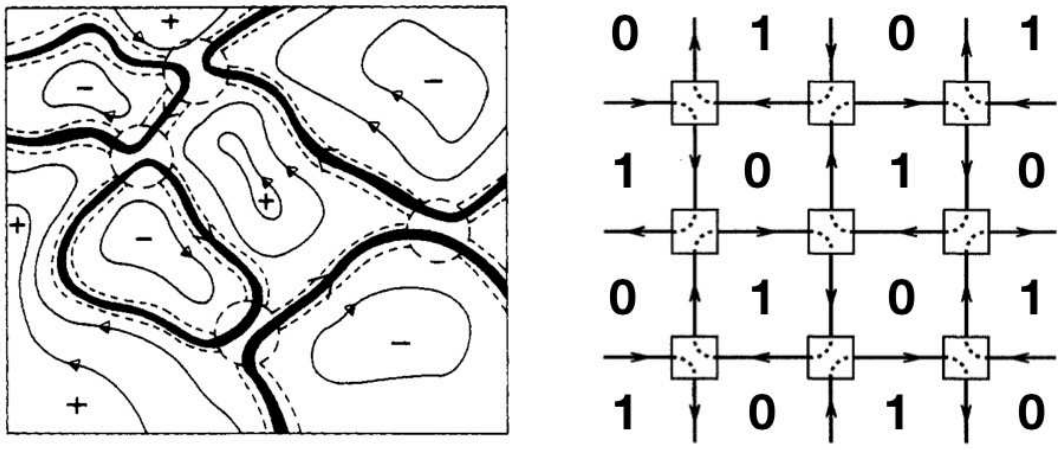


Figure 2.12: The Chalker-Coddington network model of quantum percolation. Left: Contour plots of a random potential $V(x, y)$. Local maxima and minima are denoted by $+$ and $-$ symbols. From J. T. Chalker and P. D. Coddington, *J. Phys. C: Solid State Phys.* **21**, 3665 (1988). Right: Idealized network of saddle points. Dashed curves show $\gamma \to \infty$ limit of the $\mathcal{S}$-matrix scattering (see text). From D.-H. Lee, Z. Wang, and S. Kivelson, *Phys. Rev. Lett.* **70**, 4130 (1993).

$|o|^2 - |i'|^2$ which is a restatement of current conservation $|o|^2 + |o'|^2 = |i|^2 + |i'|^2$. With $r = \sin(\theta)\, e^{i\psi}$ and $t = \cos(\theta)\, e^{i\omega}$, we have

$$\mathcal{M} = \begin{pmatrix} 1 & 0 \\ 0 & -e^{i(\psi+\delta)} \end{pmatrix} \begin{pmatrix} \sec\theta & \tan\theta \\ \tan\theta & \sec\theta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -e^{-i\psi} \end{pmatrix} e^{i\omega} \tag{2.71}$$

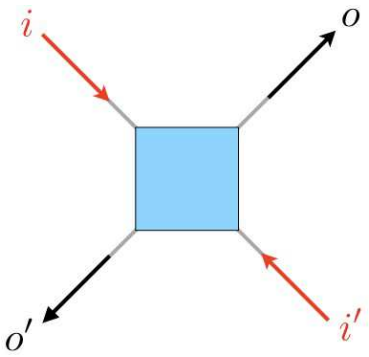
In the network model, flux amplitudes accrue random phases due to the varying lengths of the trajectories between saddles (see Eqn. 2.13), and the phases $e^{i\psi}$, $e^{i\omega}$, and $e^{i\delta}$ can be absorbed. Thus, we may take the transfer matrix at saddle point $\boldsymbol{r}$ to be

$$\mathcal{M}_{\boldsymbol{r}} = \begin{pmatrix} \sec\theta_{\boldsymbol{r}} & \tan\theta_{\boldsymbol{r}} \\ \tan\theta_{\boldsymbol{r}} & \sec\theta_{\boldsymbol{r}} \end{pmatrix} \equiv \begin{pmatrix} \cosh\gamma_{\boldsymbol{r}} & \sinh\gamma_{\boldsymbol{r}} \\ \sinh\gamma_{\boldsymbol{r}} & \cosh\gamma_{\boldsymbol{r}} \end{pmatrix} \tag{2.72}$$

with the identification $\cosh\gamma_{\boldsymbol{r}} \equiv \sec\theta_{\boldsymbol{r}}$ for $\theta_{\boldsymbol{r}} \in [0, \frac{\pi}{2}]$. While the parameter $\gamma_{\boldsymbol{r}}$ can vary from saddle to saddle, the simplest model takes $\theta_{\boldsymbol{r}} = \theta(E)$, hence $\gamma_{\boldsymbol{r}} = \gamma(E)$, to be the same energy-dependent value at all saddles. At $E = 0$, the reflection and transmission amplitudes are identical, hence $\theta(0) = \frac{\pi}{4}$ and $\gamma(0) = \ln\bigl(1 + \sqrt{2}\bigr)$.

Linear chain of saddle points

The simplest case to consider involving many saddles is that of the linear chain, depicted in Fig. 2.14. This setting is essentially that from the Mil'nikov-Sokolov argument in §2.1.6. We



$$\begin{pmatrix} o' \\ o \end{pmatrix} = \mathcal{S} \begin{pmatrix} i \\ i' \end{pmatrix}$$

$$\begin{pmatrix} o \\ i' \end{pmatrix} = \mathcal{M} \begin{pmatrix} i \\ o' \end{pmatrix}$$

Figure 2.13: Relation between scattering ($\mathcal{S}$) and transfer ($\mathcal{M}$) matrices.

have, from the figure, $\Psi_{n+1} = U_n \, \mathcal{M}_n \, \Psi_n$, with

$$\overbrace{\begin{pmatrix} i_{n+1} \\ o'_{n+1} \end{pmatrix}}^{\Psi_{n+1}} = \overbrace{\begin{pmatrix} e^{i\alpha_n} & 0 \\ 0 & e^{-i\beta_n} \end{pmatrix}}^{U_n} \overbrace{\begin{pmatrix} \cosh\gamma_n & \sinh\gamma_n \\ \sinh\gamma_n & \cosh\gamma_n \end{pmatrix}}^{\mathcal{M}_n} \overbrace{\begin{pmatrix} i_n \\ o'_n \end{pmatrix}}^{\Psi_n} \quad . \tag{2.73}$$

where α_n and β_n are the phases accrued for the right- and left-moving flux amplitudes between saddles n and $n+1$. Thus, after N such saddles, we have $\Psi_{N+1} = \mathcal{Q}_N \, \Psi_1$, where the cumulative transfer matrix $\mathcal{Q}_N$ is given by

$$\mathcal{Q}_N \ = U_N \, \mathcal{M}_N \, \cdots \, U_2 \, \mathcal{M}_2 \, U_1 \, \mathcal{M}_1 \quad . \tag{2.74}$$

Suppose we wish to calculate the disorder average $\langle \, \Psi^\dagger_{N+1} \, A \, \Psi_{N+1} \, \rangle$, where $A = a_0 + \boldsymbol{a} \cdot \boldsymbol{\sigma}$ is an arbitrary 2×2 Hermitian matrix. Clearly $\langle \, \Psi^\dagger_{N+1} \, A \, \Psi_{N+1} \, \rangle = \Psi^\dagger_1 \, \langle \mathcal{Q}^\dagger_N \, A \, \mathcal{Q}_N \rangle \, \Psi_1$, so let us compute

$$\langle \mathcal{Q}^\dagger_N \, A \, \mathcal{Q}_N \rangle = \langle \, \mathcal{M}^\dagger_1 \, U^\dagger_1 \cdots \mathcal{M}^\dagger_N \, U^\dagger_N \, A \, U_N \, \mathcal{M}_N \cdots U_1 \, \mathcal{M}_1 \, \rangle \tag{2.75}$$

We start in the middle, first averaging over the random variables α_N and β_N, which are presumed to be independent and uniformly distributed over the circle. We have

$$\left\langle \begin{pmatrix} e^{-i\alpha_N} & 0 \\ 0 & e^{i\beta_N} \end{pmatrix} \begin{pmatrix} a_0 + a_3 & a_1 - i a_2 \\ a_1 + i a_2 & a_0 - a_3 \end{pmatrix} \begin{pmatrix} e^{i\alpha_N} & 0 \\ 0 & e^{-i\beta_N} \end{pmatrix} \right\rangle = \begin{pmatrix} a_0 + a_3 & 0 \\ 0 & a_0 - a_3 \end{pmatrix} \tag{2.76}$$

and

$$\begin{pmatrix} \cosh\gamma_N & \sinh\gamma_N \\ \sinh\gamma_N & \cosh\gamma_N \end{pmatrix} \begin{pmatrix} a_0 + a_3 & 0 \\ 0 & a_0 - a_3 \end{pmatrix} \begin{pmatrix} \cosh\gamma_N & \sinh\gamma_N \\ \sinh\gamma_N & \cosh\gamma_N \end{pmatrix} \tag{2.77}$$

$$= \begin{pmatrix} \cosh(2\gamma_N) \, a_0 + a_3 & \sinh(2\gamma_N) \, a_0 \\ \sinh(2\gamma_N) \, a_0 & \cosh(2\gamma_N) \, a_0 - a_3 \end{pmatrix} \quad .$$

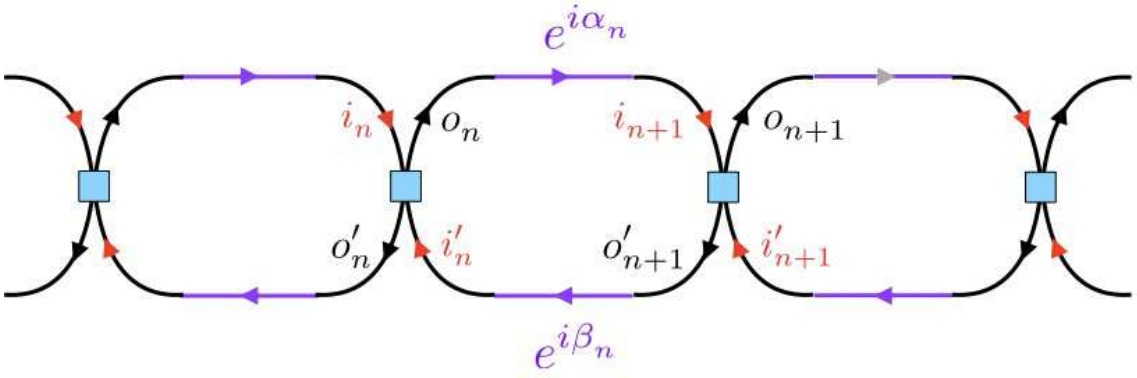


Figure 2.14: A chain of quantum saddle points.

Thus, we have the iterative rule

$$\begin{pmatrix} a'_0 \\ a'_1 \\ a'_2 \\ a'_3 \end{pmatrix} = \begin{pmatrix} \cosh(2\gamma_N)\, a_0 \\ \sinh(2\gamma_N)\, a_0 \\ 0 \\ a_3 \end{pmatrix} \tag{2.78}$$

and therefore

$$\left\langle\, \mathcal{Q}_N^\dagger \left(a_0\, 1 + \boldsymbol{a}\cdot\boldsymbol{\sigma}\right) \mathcal{Q}_N \,\right\rangle = \prod_{n=1}^{N} \cosh(2\gamma_N)\, a_0 \cdot \left(1 + \tanh(2\gamma_1)\, \sigma^1\right) + a_3\, \sigma^3 \quad . \tag{2.79}$$

Note that $\langle \mathcal{Q}_N^\dagger\, Z\, \mathcal{Q}_N \rangle = Z$, which is again the condition of current conservation. It is convenient to eliminate the σ^1 term by computing instead $\langle\, \Psi_{N+1}^\dagger\, A\, \Psi_{N+1}\,\rangle = \Phi_0^\dagger\, \langle U_0^\dagger\, \mathcal{Q}_N^\dagger\, A\, \mathcal{Q}_N\, U_0 \rangle\, \Phi_0$, where $\Phi_0^\dagger = \begin{pmatrix} o_0^* & i_0'^* \end{pmatrix}$, in which case

$$\left\langle\, U_0^\dagger\, \mathcal{Q}_N^\dagger \left(a_0\, 1 + \boldsymbol{a}\cdot\boldsymbol{\sigma}\right) \mathcal{Q}_N\, U_0 \,\right\rangle = \Lambda_N\, a_0\, 1 + a_3\, \sigma^3 \quad . \tag{2.80}$$

where $\Lambda_N \equiv \prod_{n=1}^{N} \cosh(2\gamma_n)$. With $\widetilde{\mathcal{Q}}_N \equiv \mathcal{Q}_N\, U_0$, we then have $\Psi_{N+1} = \widetilde{\mathcal{Q}}_N\, \Phi_0$, *i.e.*

$$\begin{pmatrix} i_{N+1} \\ o'_{N+1} \end{pmatrix} = \overbrace{\begin{pmatrix} 1/\mathsf{t}_N^* & -\mathsf{r}_N^*/\mathsf{t}_N^* \\ -\mathsf{r}_N\, e^{i\delta_N}/\mathsf{t}_N^* & e^{i\delta_N}/\mathsf{t}_N^* \end{pmatrix}}^{\widetilde{\mathcal{Q}}_N} \begin{pmatrix} o_0 \\ i'_0 \end{pmatrix} \quad . \tag{2.81}$$

From this we obtain

$$\begin{aligned} \left\langle |i_{N+1}|^2 \right\rangle &= \frac{\Lambda_N + 1}{2}\, |o_0|^2 + \frac{\Lambda_N - 1}{2}\, |i'_0|^2 \\ \left\langle |o'_{N+1}|^2 \right\rangle &= \frac{\Lambda_N - 1}{2}\, |o_0|^2 + \frac{\Lambda_N + 1}{2}\, |i'_0|^2 \quad , \end{aligned} \tag{2.82}$$

with $\frac{1}{2}(\Lambda_N + 1) = \langle 1/|\mathsf{t}_N|^2\rangle$. Note that terms proportional to $o_0^* \, i_0'$ do not enter on the RHS above because their coefficients $\langle \mathsf{r}/|\mathsf{t}|^2\rangle$ vanish because of the phase averaging. We identify the effective transmission and reflection coefficients by setting $o_0 = 0$ and evaluating

$$\mathsf{T}_N \equiv \frac{\langle |i_{N+1}|^2\rangle}{|o_0|^2} = \frac{2}{\Lambda_N + 1} \quad , \tag{2.83}$$

with $\mathsf{R}_N = 1 - \mathsf{T}_N$. If the $\{\gamma_n\}$ are randomly distributed, one must compute the corresponding average $\langle \Lambda_N \rangle = \langle \cosh(2\gamma)\rangle^N$ over the distribution $P(\gamma)$. The net result is that the transmission coefficient decays exponentially, as $\mathsf{T}_N \propto \exp(-N/\xi)$ with $\xi(\gamma) = 1/\ln\cosh(2\gamma)$. This is equivalent to one-dimensional Anderson localization.

Square lattice network model

Consider now the square lattice network model defined in Fig. 2.15. Each vertex is again described by a 2×2 $\mathcal{S}$-matrix as above, and the phases along the links act as 1×1 $\mathcal{S}$-matrices in the following manner:

$$\begin{aligned} o_{j+1\,,\,k} &= \exp(-i\,\alpha_{j\,,\,k})\, i'_{j\,,\,k} \\ i'_{j\,,\,k+1} &= \exp(i\,\beta_{j\,,\,k})\, o'_{j\,,\,k} \\ i_{j+2\,,\,k} &= \exp(i\,\alpha_{j+1\,,\,k})\, o'_{j+1\,,\,k} \\ o_{j+1\,,\,k+1} &= \exp(-i\beta_{j+1\,,\,k})\, i_{j+1\,,\,k} \quad , \end{aligned} \tag{2.84}$$

where $j + k$ is even, corresponding to the blue vertices in the figure. At each vertex, we have

$$\begin{pmatrix} o' \\ o \end{pmatrix} = \begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \begin{pmatrix} i \\ i' \end{pmatrix} \qquad \Longleftrightarrow \qquad \begin{pmatrix} o \\ i' \end{pmatrix} = \begin{pmatrix} \sec\theta & \tan\theta \\ \tan\theta & \sec\theta \end{pmatrix} \begin{pmatrix} i \\ o' \end{pmatrix} \quad . \tag{2.85}$$

In the γ parameterization, recall that $\cosh\gamma = \sec\theta$, *i.e.* $\gamma = \ln(\sec\theta + \tan\theta)$. When $\theta = 0$ ($\gamma = 0$) we have $o = i$ and $o' = i'$, *i.e.* perfect transmission. In Fig. 2.15, this corresponds to clockwise motion around $\nu > \frac{1}{2}$ regions, *i.e.* $E > 0$. When $\theta = \frac{\pi}{2}$ ($\gamma = \infty$), we have $o' = -i$ and $o = i'$, *i.e.* perfect reflection, corresponding to counterclockwise motion around $\nu < \frac{1}{2}$ regions.

In the numerical work by Lee, Wang, and Kivelson (LWK)[38], the scattering parameter γ was given by $\gamma_{\boldsymbol{r}} = \gamma_{\rm c} \exp(\mu - V_{\boldsymbol{r}})$, where $\gamma_{\rm c} = \ln(1 + \sqrt{2})$ is the critical value (corresponding to $\theta = \frac{\pi}{4}$), μ is a dimensionless chemical potential, and $V_{\boldsymbol{r}}$ is a dimensionless local random potential distributed uniformly on the interval $[-W, W]$. The system is arranged in a $L \times M$ rectangular lattice, with the M direction periodic. By applying the $\mathcal{S}$-matrices at each vertex and along the links, one can numerically derive an $M \times M$ transfer matrix $\mathcal{T}(L, M, \mu)$ which acts on the leftmost column of M flux amplitudes to generate the rightmost column of M flux

[38] D.-H. Lee, Z. Wang, and S. Kivelson, *Phys. Rev. Lett.* **70**, 4130 (1993).

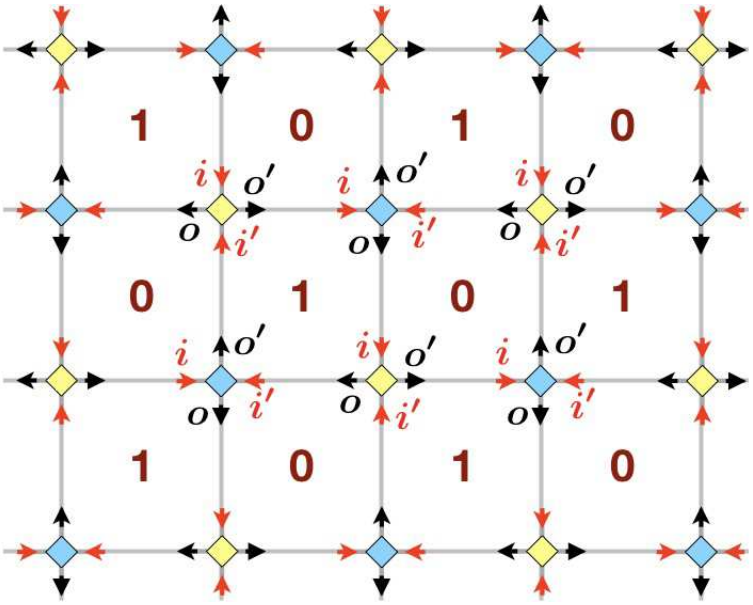


Figure 2.15: The square lattice Chalker-Coddington network model. Plaquettes marked with 0 correspond to $\nu < \frac{1}{2}$ regions, while plaquettes marked with 1 correspond to $\nu > \frac{1}{2}$ regions.

amplitudes[39], *viz.*

$$\begin{pmatrix} o'_1(L) \\ \vdots \\ o'_{M/2}(L) \\ i'_1(L) \\ \vdots \\ i'_{M/2}(L) \end{pmatrix} = \mathcal{T} \begin{pmatrix} i_1(1) \\ \vdots \\ i_{M/2}(L) \\ o_1(L) \\ \vdots \\ o_{M/2}(L) \end{pmatrix} \quad . \tag{2.86}$$

Note that if $P(\lambda) = \det(\lambda - \mathcal{T})$ is the characteristic polynomial of $\mathcal{T}$, then $\mathcal{T}^\dagger Z \mathcal{T} = Z$, where $Z = \mathsf{diag}(1, \ldots, 1, -1, \ldots, -1) = 1_{M/2} \oplus (-1)_{M/2}$ is a diagonal $M \times M$ matrix, then

$$P(\lambda) = \lambda^M \det(\lambda^{-1} - \mathcal{T}^*) / \det(\mathcal{T}^*) \quad , \tag{2.87}$$

which establishes that the roots of $P(\lambda)$ come in pairs $(\lambda_j, 1/\lambda_j^*)$, where without loss of generality we may assume $|\lambda_j| \geq 1$ for $j \in \{1, \ldots, \frac{1}{2}M\}$. If the eigenvalues $\lambda_j(L, M, \mu)$ of $\mathcal{T}(L, M, \mu)$ are ordered such that $1 \leq |\lambda_1| \leq \cdots \leq |\lambda_{M/2}|$, then we define the j^{th} *localization length* $\xi_j(M, \mu)$ as the (inverse of the) positive real quantity

$$\xi_j^{-1}(M, \mu) \equiv \lim_{L \to \infty} \frac{1}{L} \ln \big| \lambda_j(L, M, \mu) \big| \quad . \tag{2.88}$$

As defined, we have $M/2$ localization lengths, ordered according to $\xi_1(M, \mu) \geq \cdots \geq \xi_{M/2}(M, \mu)$. We are interested in the largest localization length, *i.e.* for $j = 1$. In the limit $W \to \infty$, the pa-

[39] We assume L and M are even.

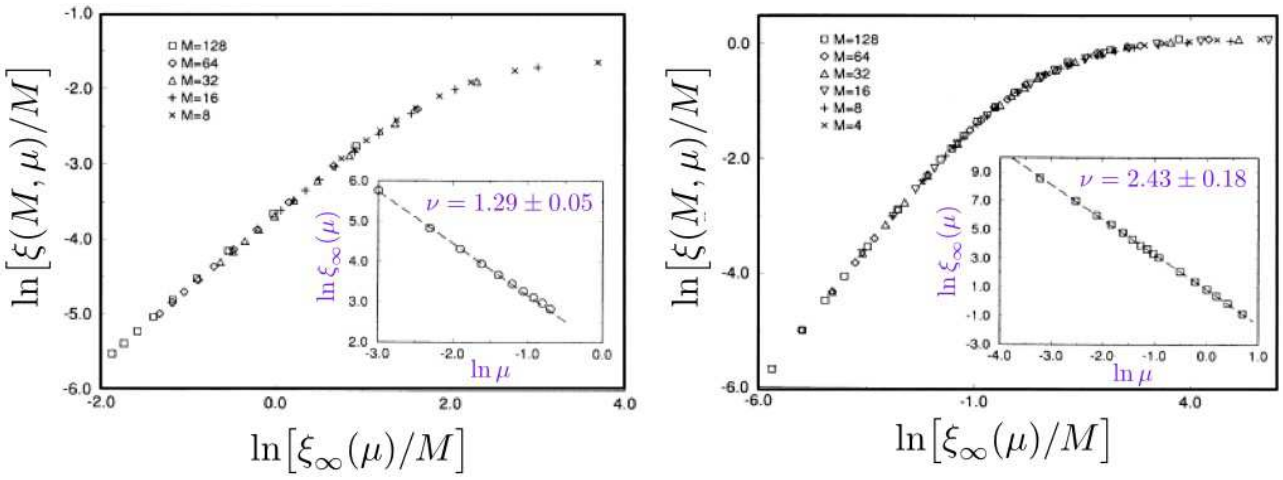


Figure 2.16: Scaling determination of exponent ν for classical and quantum 2D percolation. From D.-H. Lee, Z. Wang, and S. Kivelson, *Phys. Rev. Lett.* **70**, 4130 (1993).

rameter γ is finite with vanishing probability, and the classical percolation limit is recovered. For quantum percolation, LWK set $W = 1$. LWK examined systems with $L = 2 \times 10^5$ and with M ranging from 8 to 128. Finite size scaling theory says that when $L \gg M$, one can write

$$\xi(M,\mu) = M F(\xi_\infty(\mu)/M) \quad , \tag{2.89}$$

where F is a scaling function and where $\xi_\infty(\mu)$ is the thermodynamic localization length. To obtain $\xi_\infty(\mu)$, one plots $\ln\left[\xi(M,\mu)/M\right]$ *versus* $\ln(1/M)$ for different values of μ and then shifts the curve for each μ by an amount such that the curves all overlap. The amount of the shift is then identified with $\ln \xi_\infty(\mu)$. This method yields both the thermodynamic localization length $\xi_\infty(\mu)$ as well as the scaling function F.

2.2.4 Tight-binding and other models of the disordered Landau Level

Other numerical investigations of the IQH transition have utilized a Hamiltonian approach, where when projected on the $n^{\rm th}$ Landau level the Hamiltonian matrix elements are

$$H_{nk_1,nk_2} = (n + \tfrac{1}{2})\hbar\omega_{\rm c}\,\delta_{k_1,k_2} + \langle\, n,k_1 \,|\, V \,|\, n,k_2 \,\rangle \quad , \tag{2.90}$$

where $V(\boldsymbol{r})$ is the random potential and $k_{1,2}$ are the y-wavevectors in the Landau strip basis. Because the Landau strip spacing is $\Delta x = 2\pi\ell^2/L_y$, in the limit $L_y \to \infty$, the random potential couples many different strip wavefunctions, leading to a long-ranged one-dimensional hopping problem in the k_y basis, where, please recall, $\exp(ik_yL_y) = 1$. Another lattice model is the disordered Hofstadter model, whose real space matrix elements are given by

$$H_{\boldsymbol{r},\boldsymbol{r}'} = -t\, e^{iA_{\boldsymbol{r}\boldsymbol{r}'}} + W_{\boldsymbol{r}}\,\delta_{\boldsymbol{r},\boldsymbol{r}'} \quad , \tag{2.91}$$

with $A_{r'r} = -A_{rr'}$, and where the directed (counterclockwise) sum of the gauge potential $A_{rr'}$ along nearest neighbor links $\langle rr' \rangle$ around a plaquette p gives the dimensionless flux ϕ_p. Typically one then computes the lattice Green's function for this model,

$$G_{r,r'}(E) = \langle\, r \,|\, \frac{1}{E + i\epsilon - H} \,|\, r' \,\rangle \tag{2.92}$$

as a function of energy E and fixed dimensionless ratio W/t, where W is the width of the distribution of the $\{W_r\}$. On a strip of length L and width M, the longest localization length is given by the expression

$$\xi^{-1}(M, E) = -\lim_{L\to\infty} \lim_{\epsilon\to 0} \frac{1}{2L} \left\langle \ln \sum_{i,j=1}^{M} \left| G_{1i,Lj}(E) \right|^2 \right\rangle \quad , \tag{2.93}$$

where $i, j \in \{1, \ldots, M\}$ label the transverse coordinate[40]. The approach of using Green's functions to numerically compute localization lengths was pioneered by MacKinnon and Kramer in the early 1980s[41]. The scaling theory of the IQH transition has been reviewed by B. Huckestein[42]. A computational advantage of this method over exact diagonalization is that the Green's function can be computed recursively. Very recently, Puschmann *et al.*[43] obtained $\nu = 2.58(3)$ in studies of the disordered square lattice Hofstadter model, consistent with Slevin and Ohtsuki's network model result $\nu = 2.593 \pm 0.006$[44]. Note that this rules out the Mil'nikov-Sokolov result $\nu_{\rm QU} = \frac{7}{3}$, which is in fact closer to the experimentally determined value of ν.

Zhu, Wu, Bhatt, and Wan[45], building on earlier work of Huo and Bhatt[46], investigated a disordered square lattice Hofstadter model, equivalent to Eqn. 2.91, with $\phi = \frac{2\pi}{3}$ per structural unit cell. In the absence of disorder, this yields three Landau subbands $|\psi_n(\boldsymbol{\theta})\rangle$, with Chern numbers $+1$, 0, and -1, for $n = 1, 2$, and 3, respectively. They then considered the truncated model where the disorder potential W_r is projected onto the $n = 1$ subband of the disorder-free model, *i.e.*

$$H(\boldsymbol{\theta}, \boldsymbol{\theta}') = \sum_{\boldsymbol{\theta},\boldsymbol{\theta}'} |\, \psi_1(\boldsymbol{\theta}) \,\rangle\langle\, \psi_1(\boldsymbol{\theta}) \,|\, W_{\boldsymbol{r}} \,|\, \psi_1(\boldsymbol{\theta}') \,\rangle\langle\, \psi_1(\boldsymbol{\theta}') \,| \quad , \tag{2.94}$$

where $\theta_{1,2} = (2\pi j_{1,2} + \zeta)_{1,2})/N_{1,2}$, where $j_{1,2} \in \{1, \ldots, N_{1,2}\}$ and where $\alpha_{1,2} \in [0, 2\pi]$. Note we could also denote $H(\boldsymbol{\theta}, \boldsymbol{\theta}') = H_{\boldsymbol{j},\boldsymbol{j}'}(\boldsymbol{\zeta})$. There are then $N_1 N_2$ eigenstates $|\, \varphi_l(\boldsymbol{\zeta}) \,\rangle$ for each pair of boundary phases $(\zeta)_1, \zeta)_2)$, and the Chern numbers are given by

$$C_l = \frac{i}{2\pi} \int_{\mathsf{T}^2} d^2\zeta\, \epsilon_{\alpha\beta} \left\langle \frac{\partial \varphi_l}{\partial \zeta^\alpha} \,\middle|\, \frac{\partial \varphi_l}{\partial \zeta^\beta} \right\rangle \tag{2.95}$$

[40] In systems of higher dimension $d > 2$ and of size $M \times \cdots \times M \times L$, one writes $i \to \boldsymbol{r}_\perp$ and $j \to \boldsymbol{r}'_\perp$, and the sum is over all $M^{2(d-1)}$ pairs $(\boldsymbol{r}_\perp, \boldsymbol{r}'_\perp)$.

[41] See A. MacKinnon and B. Kramer, *Phys. Rev. Lett.* **47**, 1546 (1981). For application to the IQHE, see B. Huckestein and B. Kramer, *Phys. Rev. Lett.* **64**, 1437 (1990).

[42] B. Huckestein, *Rev. Mod. Phys.* **67**, 357 (1995).

[43] M. Puschmann, P. Cain, M. Schreiber, and T. Vojta, *Phys. Rev. B* **99**, 121301(R) (2019).

[44] K. Slevin and T. Ohtsuki, *Phys. Rev. B* **80**, 041304 (2009)

[45] Q. Zhu, P. Wu, R. N. Bhatt, and X. Wan, *Phys. Rev. B* **99**, 024205 (2019).

[46] Y. Huo and R. N. Bhatt, *Phys. Rev. Lett.* **68**, 1375 (1992).

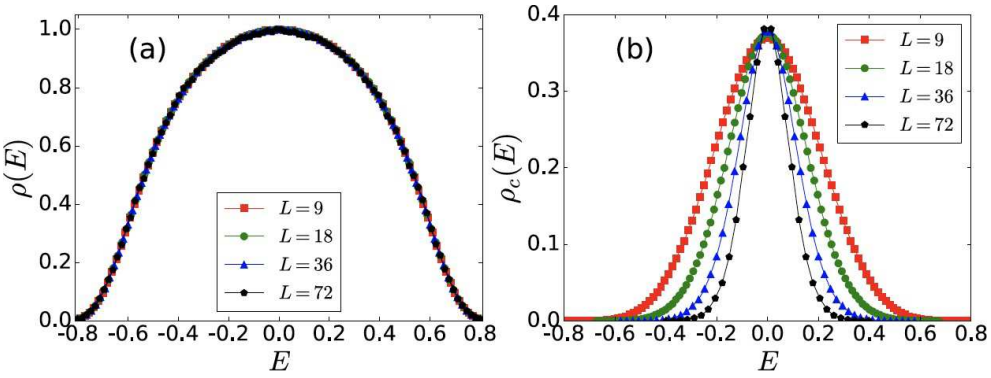


Figure 2.17: Disorder-averaged DOS and DOCS (density of conducting states). From Q. Zhu, P. Wu, R. N. Bhatt, and X. Wan, *Phys. Rev. B* **99**, 024205 (2019).

where the integral is over the torus $\boldsymbol{\zeta} \in [0, 2\pi] \times [0, 2\pi]$. Zhu *et al.* computed both the disorder-averaged density of states $\rho(E)$ as well as the disorder-averaged density of conducting states $\rho_{\rm c}(E)$, defined as

$$\begin{aligned} \rho(E) &= \left\langle \frac{1}{N_1 N_2} \sum_{l=1}^{N_1 N_2} \delta(E - E_l) \right\rangle \\ \rho(E) &= \left\langle \frac{1}{N_1 N_2} \sum_{l=1}^{N_1 N_2} (1 - \delta_{C_l,0})\, \delta(E - E_l) \right\rangle \quad , \end{aligned} \tag{2.96}$$

where the second sum includes contributions only from states of nonzero Chern number. Their results are shown in Fig. 2.17. As the linear system size $N_1 = N_2 \equiv L$ is increased, the width of $\rho(E)$ remains unchanged, but that of $\rho_{\rm c}(E)$ narrows, indicating that in the thermodynamic limit a sub-extensive number of states carry nonzero Chern number. A scaling *Ansatz*, with $N_{\rm c} = a\, N_\phi^{-1/(2\nu)}$, where $N_\phi = N_1 N_2\, \phi = \frac{2\pi}{3} L^2$, was analyzed, and good agreement was found with $\nu = 2.49 \pm 0.01$[47]. The form of the scaling *Ansatz* is inspired by the fact that one expects that for system of linear size $L \propto \sqrt{N_\phi}$, states with $\xi(E) > L$ are conducting. The number of these states scales as $N_{\rm c} \sim L^2 \rho(E_{\rm c})\, |E - E_{\rm c}|$ with $\xi(E) = L$, hence $|E| \propto L^{-1/\nu}$, and $E_{\rm c} = 0$ for symmetrically distributed disorder. Thus, $N_{\rm c} \propto L^{2-1/\nu} \propto N_\phi^{1-1/(2\nu)}$.

2.2.5 Real Space Renormalization

Symmetry dictates that a square lattice Chalker-Coddington (CC) network model composed of identical scatterers with $\gamma_r = \gamma$ should exhibit a quantum critical point at $\gamma = \gamma_{\rm c} = \ln(1+\sqrt{2})$, *i.e.* at $T = R = \frac{1}{2}$, where T and R are the transmission and reflection coefficients for each individual

[47] The authors also investigated corrections to scaling, with $N_{\rm c}/N_\phi = a(1 + bN_\phi^{-\gamma/2})N_\phi^{-1/(2\nu)}$, which gave better agreement, with $\nu = 2.480 \pm 0.005$.

vertex. Consider now the case where $T = \mathrm{sech}\,\gamma$ at each site is chosen from a distribution $P(T)$, and consider the problem of transmission through an $L \times L$ lattice of saddle point vertices. One may (at least numerically) compute the distribution $P_L(T)$ of transmission (defined, say, from left to right) across this system, averaging over all the link phases. In the limit $L \to \infty$, one expects two stable distributions, given by $P_\infty(T) = \delta(T)$ and $P_\infty(T) = \delta(1-T) = \delta(R)$, corresponding to bulk localized phases with $T = 0$ and $T = 1$, respectively. The word "stable" in this context alludes to a notional renormalization group (RG) flow. As we have discussed above in §2.2.2, applications of RG techniques to the replica field theory of the IQHE have been intractable due to difficulties associated with the topological term and the $n \to 0$ replica limit. Here we describe a real space RG approach to the IQH transition based on the network model, following D. P. Arovas, M. Janssen, and B. Shapiro, *Phys. Rev. B* **56**, 1751 (1997), henceforth denoted as ABS97[48]. Real space RG (RSRG) schemes have the virtue of being easily implemented and physically appealing, but suffer from being completely uncontrolled and not providing any systematic way to calculate critical properties with increasing accuracy (such as going to more loops in diagrammatic field theory calculations)[49]. As applied to the CC network model, the RSRG approach obtains the distribution $P_{bL}(T)$ for a larger system of linear dimension bL in terms of $P_L(T)$. This functional relation may be represented in terms of a set of parameters $\{X_i(L)\}$ which characterize the distribution $P_L(T)$, such as the coefficients in a Chebyshev or Legendre polynomial expansion of $P_L(T)$ in the variable $x = 2T - 1$. Thus, one has

$$X_i(bL) = F_i\big(\{X_j(L)\}; b\big) \quad . \tag{2.97}$$

The fixed point distribution is then characterized by $\{X_i^*\}$, where $X_i^* = F_i\big(\{X_j^*\}; b\big)$, and the eigenvalues $\{\lambda_a\}$ of the matrix $R_{ij} = \big(\partial F_i/\partial X_j\big)_{\boldsymbol{X}^*}$ determine the relevance of the corresponding eigenvectors, which are the *scaling variables*. The positive eigenvalues define a set of critical exponents, $y_a = \ln \lambda_a / \ln b$ and the localization length exponent $\nu = \ln b / \ln \lambda_{\rm min}$ corresponds to the smallest eigenvalue $\lambda_{\rm min}$. The β-functions are defined to be

$$\beta_i\big(\{X_j\}\big) = \frac{\partial X_i}{\partial \ln L} = \left.\frac{\partial F_i}{\partial b}\right|_{b=1} \quad . \tag{2.98}$$

In order to implement this program exactly, one would need to compute the distribution $P_L(T)$ for finite networks of arbitrary size. This is an intractable problem for even modest $L \sim 10$, hence one must resort to some approximation scheme, which is the source of all troubles with the RSRG approach. Here we will briefly describe two such approximation schemes: (i) Migdal-Kadanoff (MK) decimation, and (ii) hierarchical lattice constructions. Both allow for a recursive implementation of the RSRG program, using only simple numerical computation.

In Fig. 2.13, we saw how the scattering matrix $\mathcal{S}$, which gives the linear relation between incoming flux amplitudes (i, i') and outgoing flux amplitudes (o, o'), can be recast as a transfer

[48] See also A. G. Galstyan and M. E. Raikh, *Phys. Rev. B* **56**, 1422 (1997).

[49] Truth be told, with the trivial exception of $d = 1$, RSRG schemes generally yield poor results for critical exponents.

matrix $\mathcal{M}$, relating data (i, o') on the left of the vertex to data (o, i') on the right. Similarly, one can define a transfer matrix $\mathcal{N}$ which relates data (i, o) above the vertex to data (o', i') below the vertex. Thus,

$$\begin{pmatrix} o' \\ o \end{pmatrix} = \mathcal{S} \begin{pmatrix} i \\ i' \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} o \\ i' \end{pmatrix} = \mathcal{M} \begin{pmatrix} i \\ o' \end{pmatrix} \quad \Leftrightarrow \quad \begin{pmatrix} o' \\ i' \end{pmatrix} = \mathcal{N} \begin{pmatrix} i \\ o \end{pmatrix} \quad . \tag{2.99}$$

In particular,

$$\mathcal{S} = \begin{pmatrix} -\sin\theta & \cos\theta \\ \cos\theta & \sin\theta \end{pmatrix} \quad \Leftrightarrow \quad \mathcal{M} = \begin{pmatrix} \sec\theta & \tan\theta \\ \tan\theta & \sec\theta \end{pmatrix} \quad \Leftrightarrow \quad \mathcal{N} = \begin{pmatrix} -\csc\theta & \operatorname{ctn}\theta \\ -\operatorname{ctn}\theta & \csc\theta \end{pmatrix} \quad . \tag{2.100}$$

Note $\mathcal{S}^\dagger \mathcal{S} = 1$ and $\mathcal{M}^\dagger Z \mathcal{M} = \mathcal{N}^\dagger Z \mathcal{N} = Z$. Now consider the combination of two scatterers in series, as depicted in the left panel of Fig. 2.18. The combined transfer matrix is given by

$$\mathcal{M}' = \mathcal{M}_2 \, U \, \mathcal{M}_1 = \begin{pmatrix} \sec\theta_2 & \tan\theta_2 \\ \tan\theta_2 & \sec\theta_2 \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\beta} \end{pmatrix} \begin{pmatrix} \sec\theta_1 & \tan\theta_1 \\ \tan\theta_1 & \sec\theta_1 \end{pmatrix} \quad . \tag{2.101}$$

Computing $\sec^2\theta' = |\mathcal{M}'_{1,1}|^2 = 1/T'$, we obtain the transmission coefficient

$$\frac{1}{T'} = \frac{1 + 2\sqrt{R_1 R_2}\,\cos(\alpha+\beta) + R_1 R_2}{T_1 T_2} \quad . \tag{2.102}$$

Averaging $\ln T$ over the angle $\phi \equiv \alpha + \beta$, we obtain $\langle \ln T' \rangle = 2 \langle \ln T \rangle$, and for b scatterers in series,

$$\langle \ln T' \rangle = b \, \langle \ln T \rangle \quad . \tag{2.103}$$

Thus $\ln T$ is driven to increasingly negative values under iteration, which is the essence of one-dimensional localization.

Equivalently, though, we may construct the transfer matrix from "top to bottom", in which case

$$\mathcal{N}' = \mathcal{N}_2 \, U \, \mathcal{N}_1 = \begin{pmatrix} -\csc\theta_2 & \operatorname{ctn}\theta_2 \\ -\operatorname{ctn}\theta_2 & \csc\theta_2 \end{pmatrix} \begin{pmatrix} e^{i\alpha} & 0 \\ 0 & e^{-i\beta} \end{pmatrix} \begin{pmatrix} -\csc\theta_1 & \operatorname{ctn}\theta_1 \\ -\operatorname{ctn}\theta_1 & \csc\theta_1 \end{pmatrix} \quad . \tag{2.104}$$

But now $\csc^2\theta' = |\mathcal{N}'_{1,1}|^2 = 1/R'$, and we have

$$\frac{1}{R'} = \frac{1 + 2\sqrt{T_1 T_2}\,\cos(\alpha+\beta) + T_1 T_2}{R_1 R_2} \quad . \tag{2.105}$$

This yields

$$\langle \ln R' \rangle = b \, \langle \ln R \rangle \tag{2.106}$$

for b scatterers in parallel. In this case it is the reflection amplitude which is driven to zero! In the network model, both series as well as parallel propagation are present, and in a sense it is the competition between these two one-dimensional localization mechanisms which gives rise to a quantum critical point describing the IQH transition. Tractably separating the series and parallel processes, however, can only be implemented in approximation schemes such as MK or hierarchical lattice constructions.

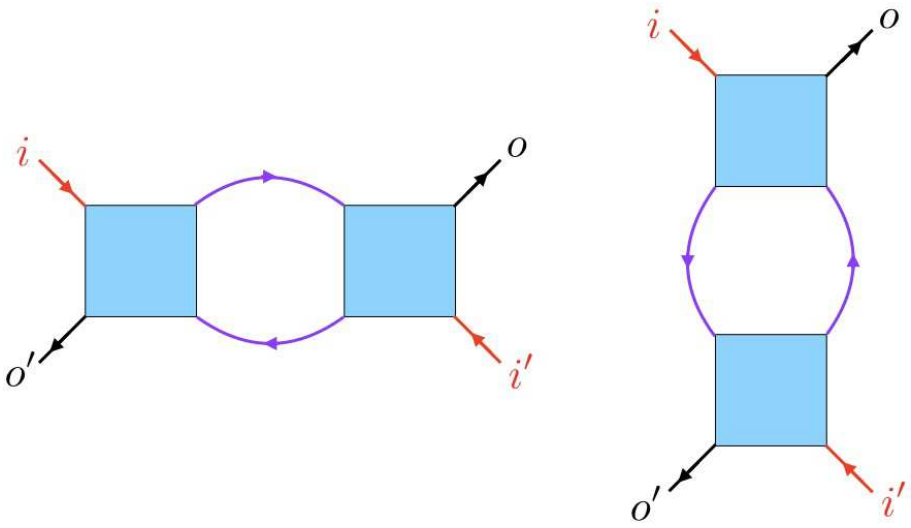


Figure 2.18: Series (left) and parallel (right) two-channel quantum scatterers.

Migdal-Kadanoff method

The Migdal-Kadanoff decimation scheme involves "bond-shifting" and is represented graphically in Fig. 2.19. To understand how this leads to critical behavior, consider the behavior of the *typical* transmission coefficient $\mathsf{T} \equiv \exp\langle \ln T \rangle$. From the above series and parallel computations, we have $\mathsf{T}_b = \mathsf{T}^b$ for series and $\mathsf{R}_b = \mathsf{R}^b$ for parallel transmission. The renormalized vertex for $b = 2$ after bond-shifting is shown in the upper right panel of Fig. 2.19, and corresponds to parallel transmission between two series pairs of the original vertices. For arbitrary b, then,

$$\mathsf{T}' = 1 - (1 - \mathsf{T}^b)^b \equiv f(\mathsf{T}; b) \quad . \tag{2.107}$$

For $b = 2$, we have $f(\mathsf{T}; b = 2) = 2\mathsf{T}^2 - \mathsf{T}^4$. This map has two stable fixed points at $\mathsf{T}^* = 0$ and 1, and an unstable fixed point at $\mathsf{T}^* = \frac{1}{2}(\sqrt{5} - 1) \simeq 0.618$. Linearizing about the unstable fixed point, we obtain the eigenvalue

$$\lambda = \left. \frac{\partial f(\mathsf{T}; b = 2)}{\partial \mathsf{T}} \right|_{\mathsf{T}^*} = 6 - 2\sqrt{5} \quad , \tag{2.108}$$

corresponding to a localization length exponent $\nu = \ln b / \ln \lambda \simeq 1.635$. Note that $\mathsf{T}^* \neq \frac{1}{2}$ because the order of the bond shifting matters[50]. Choosing instead parallel followed by series propagation, rather than series followed by parallel, the roles of T and R are reversed[51].

[50] Thus, the composite vertex in Fig. 2.19 is not invariant under 90° rotations.

[51] The RG equation in Eqn. 2.108 and its $\mathsf{T} \leftrightarrow \mathsf{R}$ counterpart coincide with the two RSRG equations obtained in the MK approach to classical bond percolation. There, the bond occupation probability p plays the role of our T, and the MK bond-shifting which in our model leads to a series or parallel composition of quantum resistors corresponds to multiplication of bond occupation probabilities. See S. R. Kirkpatrick, *Phys. Rev. B* **15**, 1533 (1977).

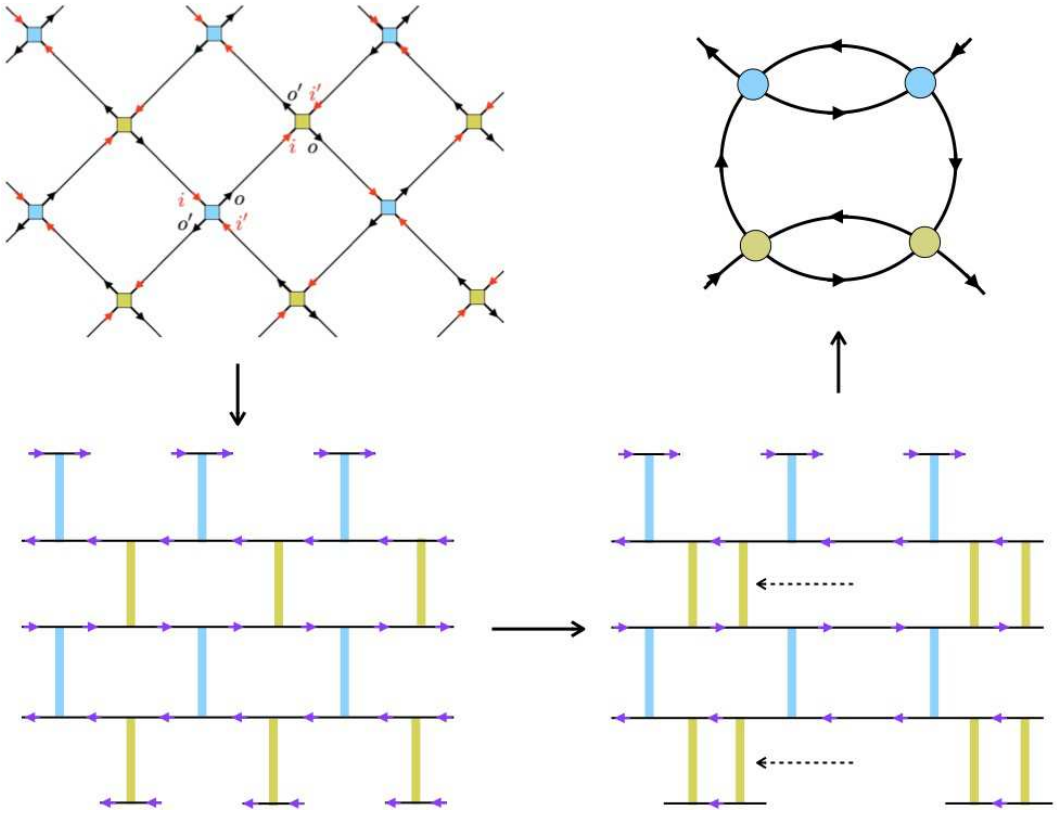


Figure 2.19: Migdal-Kadanoff decimation of the Chalker-Coddington network model. In the first stage of the decimation, the CC network is represented as a "brick lattice". A $b = 2$ bond-shifting process effectively replaces each single vertex with a group of two vertices in series. In the second stage of the decimation, the brick lattice is represented in the orthogonal (y) direction. Bond-shifting then replaces each composite vertex from stage one with a new composite vertex arising first from serial and then parallel propagation. The resulting composite vertex for $b = 2$ is shown in the upper right portion of the figure.

If we set $b = 1 + \zeta$ with $\zeta \downarrow 0$, we obtain the "infinitesimal" MK transformation,

$$\mathsf{T}' = \mathsf{T} + \zeta\big[\mathsf{T}\ln\mathsf{T} - (1-\mathsf{T})\ln(1-\mathsf{T})\big] + \mathcal{O}(\zeta^2) \quad , \tag{2.109}$$

The infinitesimal MK transformation again has fixed points at $\mathsf{T} = 0$ and $\mathsf{T} = 1$, and its unstable fixed point lies at the symmetric value $\mathsf{T} = \frac{1}{2}$, with eigenvalue $\lambda = 2(1 - \ln 2)\zeta$, corresponding to $\nu = \big[2(1 - \ln 2)\big]^{-1} \simeq 1.629$. The β-function is

$$\beta(\mathsf{T}) = \left.\frac{\partial f(\mathsf{T}; 1+\zeta)}{\partial\zeta}\right|_{\zeta=0} = \mathsf{T}\ln\mathsf{T} - (1-\mathsf{T})\ln(1-\mathsf{T}) \quad . \tag{2.110}$$

Note that $\beta(\mathsf{T}^*) = 0$ vanishes at the fixed points, where there is no RSRG flow.

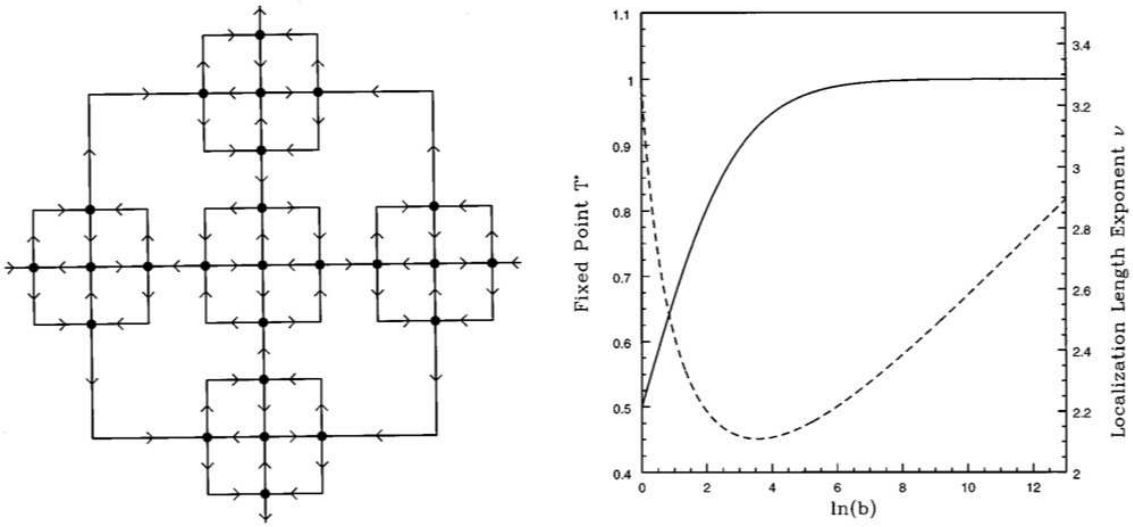


Figure 2.20: Left: A hierarchical lattice at the second stage of construction. Right: Results for a hierarchical lattice construction generalizing that in the left panel, but when the central site is replaced by a unit S-matrix. From D. P. Arovas, M. Janssen, and B. Shapiro, *Phys. Rev. B* **56**, 1751 (1997).

Hierarchical lattice constructions

A related approach to the problem involves the artifice of hierarchical lattices, which are systems of noninteger Hausdorff dimension. Consider, for example, the system depicted in the left panel of Fig. 2.20. The elementary motif is a group of $V = 5$ vertices chosen from a 3×3 ($M = 3$) group of sites, where the scatterers at the four corners are removed. This configuration is arranged into a repeating structure at ever larger length scales. Repeating this process n times results in a hierarchical structure with V^n vertices contained in a square of side length M^n. The Hausdorff dimension is t$D_{\rm H} = \ln V / \ln M$, hence $D_{\rm H}(V = 5, M = 3) \simeq 1.465$; generalizations to $V(M) = \frac{1}{2}(M^2 + 1)$ can be constructed, for which the limiting Hausdorff dimension is $\lim_{M\to\infty} D_{\rm H}(M) = 2$. To recover the previous MK scheme, replace the central scatterer with one for which $T = 1$, resulting in the four site composite vertex of Fig. 2.19. Results from the hierarchical lattice construction are shown in the right panel of Fig. 2.20.

If one replaces the central scatterer in the left panel of Fig. 2.20 with one for which $S = 1$ (perfect transmission), one recovers the four-site scattering unit of Fig. 2.19. One can use this as the fundamental unit of a hierarchical construction, and the results differ from those of the previous section only in that the linear dimension is taken to be $M = 2b - 1$ rather than $M = b$. The Hausdorff dimension of the hierarchical lattice is then $D_{\rm H} = \ln(b^2)/\ln M = 2\ln b/\ln(2b-1)$ ($= \ln 4/\ln 3 = 1.2619$ for the case $b = 2$), whereas the bond-shifted MK lattice is fully two-dimensional. The correlation length exponent ν is accordingly different, and given by $\nu = \ln(2b-1)/\ln\lambda$. For $b = 2$ one has $\nu = \ln 3/\ln(6 - 2\sqrt{5}) = 2.592$, which is shockingly (and surely fortuitously) close to the current best numerical value $\nu = 2.58$. Generalizations to larger b are

straightforward and results are shown in the right panel of Fig. 2.20 (see ABS97 for details).

2.2.6 Spin-orbit coupling

The microscopic Hamiltonian for a single electron in a potential $V(\boldsymbol{r})$ and magnetic field $\boldsymbol{B}$ is

$$H = \frac{\boldsymbol{\pi}^2}{2m_{\rm e}} + V(\boldsymbol{r}) + \frac{e\hbar}{2m_{\rm e}c}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} + \frac{\hbar}{4m_{\rm e}^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\boldsymbol{\pi} + \frac{\hbar^2}{8m_{\rm e}^2c^2}\boldsymbol{\nabla}^2V + \frac{(\boldsymbol{\pi}^2)^2}{8m_{\rm e}^3c^2} + \ldots\ , \tag{2.111}$$

where $\boldsymbol{\pi} = \boldsymbol{p} + \frac{e}{c}\boldsymbol{A}$. Where did this come from? From the Dirac equation,

$$i\hbar\frac{\partial\Psi}{\partial t} = \begin{pmatrix} m_{\rm e}c^2 + V & c\,\boldsymbol{\sigma}\cdot\boldsymbol{\pi} \\ c\,\boldsymbol{\sigma}\cdot\boldsymbol{\pi} & -m_{\rm e}c^2 + V \end{pmatrix}\Psi = E\Psi\ . \tag{2.112}$$

The wavefunction Ψ is a four-component Dirac spinor. Since $m_{\rm e}c^2$ is the largest term for our applications, the upper two components of Ψ are essentially the positive energy components. However, the Dirac Hamiltonian mixes the upper two and lower two components of Ψ. One can 'unmix' them by making a canonical transformation, $H \longrightarrow H' \equiv e^{iS}\,H\,e^{-iS}$, where S is Hermitian, to render H' block diagonal. With $E = m_{\rm e}c^2 + \varepsilon$, the effective Hamiltonian is given by (2.111). This is known as the Foldy-Wouthuysen transformation, the details of which may be found in many standard books on relativistic quantum mechanics and quantum field theory (*e.g.* Bjorken and Drell, Itzykson and Zuber, *etc.*). Note that the Dirac equation leads to $g = 2$. If we go beyond "tree level" and allow for radiative corrections within QED, we obtain a perturbative expansion, $g = 2 + \frac{\alpha}{\pi} + \mathcal{O}(\alpha^2)$, where $\alpha = e^2/\hbar c \approx 1/137$ is the fine structure constant.[52] There are two terms in (2.111) which involve the electron's spin[53]:

$$\begin{aligned} \text{Zeeman interaction}\ :\ H_{\rm Z} &= \frac{ge\hbar}{4m_{\rm e}c}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} \\ \text{Spin-orbit interaction}\ :\ H_{\rm SO} &= \frac{\hbar^2}{4m_{\rm e}^2c^2}\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V\times\left(\boldsymbol{k} + \tfrac{e}{\hbar c}\boldsymbol{A}\right)\quad . \end{aligned} \tag{2.113}$$

We define $\lambda_0 \equiv \hbar^2/4m_{\rm e}c^2 = 3.7\times10^{-6}\,\text{Å}^2$ to be the vacuum SO coupling parameter.

In crystalline solids, spin-orbit effects can be profound for large Z ions[54]. For crystalline GaAs, as well as for Si and Ge, near the Γ point in the Brillouin zone, the antibonding conduction band s-orbitals are split by the band gap Δ from the bonding valence band p-orbitals[55]. Including

[52] Note that with $\mu_{\rm n} = e\hbar/2m_{\rm p}c$ for the nuclear magneton, $g_{\rm p} = 2.793$ and $g_{\rm n} = -1.913$. These results immediately suggest that there is composite structure to the nucleons, *i.e.* quarks.

[53] The numerical value for $\mu_{\rm B}$ is $\mu_{\rm B} = e\hbar/2mc = 5.788\times10^{-9}\,\text{eV/G}$, hence $\mu_{\rm B}/k_{\rm B} = 6.717\times10^{-5}\,\text{K/G}$.

[54] For a thorough discussion of spin-orbit effects in solids, see R. Winkler, *Spin-Orbit Coupling Effects in Two-Dimensional Electron and Hole Systems* (Springer, 2003).

[55] In Si and Ge, the conduction band minimum at Γ is not the lowest energy point in the conduction band.

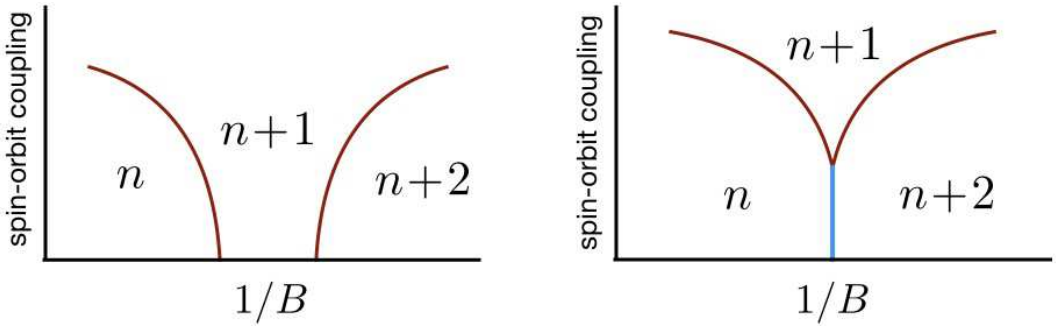


Figure 2.21: Possible phase diagrams for the disordered Landau level with spin-orbit coupling. Left: Two distinct transitions, each with $\kappa = 1/\nu z \simeq 0.42$. Right: One merged transition at sufficiently weak SO disorder scattering, with $\kappa \simeq 0.21$.

electron spin, these amount to eight states. When spin-orbit effects are included, the six valence band orbitals rearrange themselves into light and heavy hole bands which are degenerate at Γ, and a lower energy *split-off hole* band, where the splitting $\Delta_{\rm so}$ is due to spin-orbit effects. In a crystalline energy band and in the presence of a disorder potential $V_{\rm imp}(\boldsymbol{r})$ arising from impurities, the effective Hamiltonian is

$$H_{\rm eff} = \frac{\hbar^2}{2m^*}\big(\boldsymbol{k} + \tfrac{e}{\hbar c}\boldsymbol{A}\big)^2 + V_{\rm imp}(\boldsymbol{r}) + \frac{g^* e\hbar}{4m_{\rm e}c}\,\boldsymbol{\sigma}\cdot\boldsymbol{B} + \lambda\,\boldsymbol{\sigma}\cdot\boldsymbol{\nabla}V_{\rm imp}\times\big(\boldsymbol{k} + \tfrac{e}{\hbar c}\boldsymbol{A}\big) \quad , \tag{2.114}$$

which is of the same form as the expression in vacuum, but where the coupling λ is now material-dependent. For the conduction band of GaAs, for example, $\lambda = 5.3\,\text{Å}^2$, which is 10^6 times larger than the vacuum value λ_0[56].

We now ask: what happens when we include spin degrees of freedom in the IQH transition? First note that the Zeeman energy splits $\uparrow$ and $\downarrow$ spin states by $\Delta_{\rm Z} = \zeta\hbar\omega_{\rm c}$, where $\zeta \equiv g^* m^*/2m_{\rm e}$. In other words, with $V_{\rm imp}(\boldsymbol{r}) = 0$ the energy eigenvalues are $E_{n\sigma} = (n + \frac{1}{2} + \frac{1}{2}\sigma\zeta)\hbar\omega_{\rm c}$. Due to the effects of band structure and confinement in quantum wells, the g-factor can vary considerably from its tree level QED value of $g = 2$. Values as high as $g^* = 60$ have been observed in InAs/AlSb quantum wells, and g can also be tuned by pressure – in some cases to $g^* = 0$.

Consider the case of a single cyclotron Landau level with $a^\dagger a = n$ and $\sigma = \pm 1$. In the absence of SO coupling, and with weak disorder $V(\boldsymbol{r}) \equiv V_{\rm imp}(\boldsymbol{r})$ coupling only to density and not to spin, there are two independent transitions. What happens when the $\uparrow$ and $\downarrow$ spin components are mixed through the SO coupling term in H? Khmelnitskii[57] argued that the extended states of overlapping Landau spin subbands should split, and network model simulations by Lee and Chalker[58] support this conclusion, and that the localization length $\xi(E)$ thus diverges at two distinct energies. Polyakov and Shklovskii[59] further argued that if the two extended state

[56] See B. I. Halperin, *2005 Boulder Summer School for Condensed Matter and Materials Physics* lecture notes.
[57] D. E. Khmelnitskii, *Helv. Phys. Acta* **65**, 164 (1992).
[58] D. K. K. Lee and J. T. Chalker, *Phys. Rev. Lett.* **72**, 1510 (1994).
[59] D. G. Polyakov and B. I. Shklovskii, *Phys. Rev. Lett.* **70**, 3796 (1993).

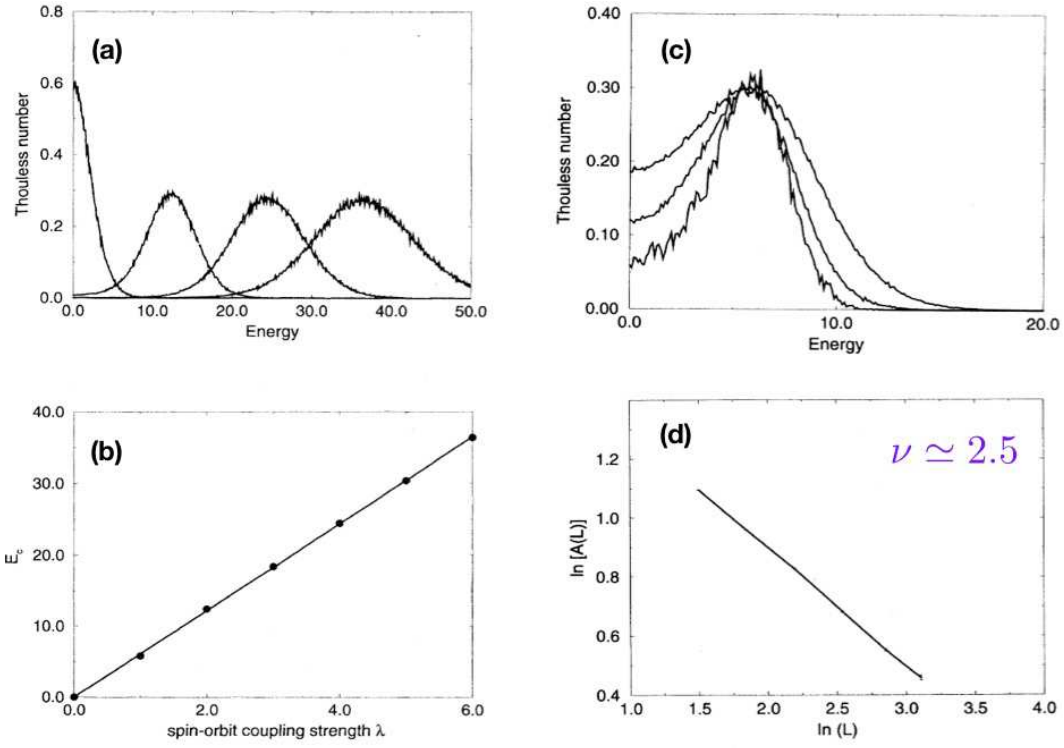


Figure 2.22: (a) Thouless number data for smooth SO scattering ($\zeta = 2$) for $N_\phi = 40$, 60, 160, and 500. (b) Log-log plot of the integrated area under the Thouless number curves $T_L(E)$ *versus* system size. The resulting slope yields $\nu \approx 2.5$. (c) Thouless number for smooth SO scattering ($\zeta = 2$) as a function of energy for coupling strengths $\lambda = 0$, 2, 4, and 6. (d) Tracking of the Thouless number peak, with $E_c(\lambda) \propto \lambda W$. From C. B. Hanna *et al.*, *Phys. Rev. B* **52**, 5221 (1995).

energies lie at $E = \pm E_c$, then the localization length should take the form

$$\xi(E) \propto \left(\frac{\Gamma^2}{|E^2 - E_c^2|} \right)^\nu \quad , \tag{2.115}$$

where Γ is the disorder broadening of the LLs, which is assumed to satisfy $\Gamma \gg E_c$. This would suggest a crossover behavior where the actual correlation length exponent ν is observed only very close to $E = \pm E_c$, and at sufficiently low temperatures. Otherwise, $\xi \sim |\Gamma/E|^{2\nu}$ and an apparent doubling of the exponent is predicted. Such an apparent doubling of the exponent for spin-degenerate peaks was reported in microwave measurements[60], where the width of the transition is observed to scale as $\Delta B \propto \omega^\gamma$, where $\gamma = 1/\nu z$ if the quantum critical point view of the transition pertains. For isolated peaks, $\gamma = 0.41 \pm 0.04$ was found, while for spin-degenerate peaks, $\gamma = 0.20 \pm 0.05$. The reader is advised to note our descriptive collocation, "*apparent* doubling" (italics for stress).

[60] L. W. Engel *et al.*, *Phys. Rev. Lett.* **71**, 2638 (1993).

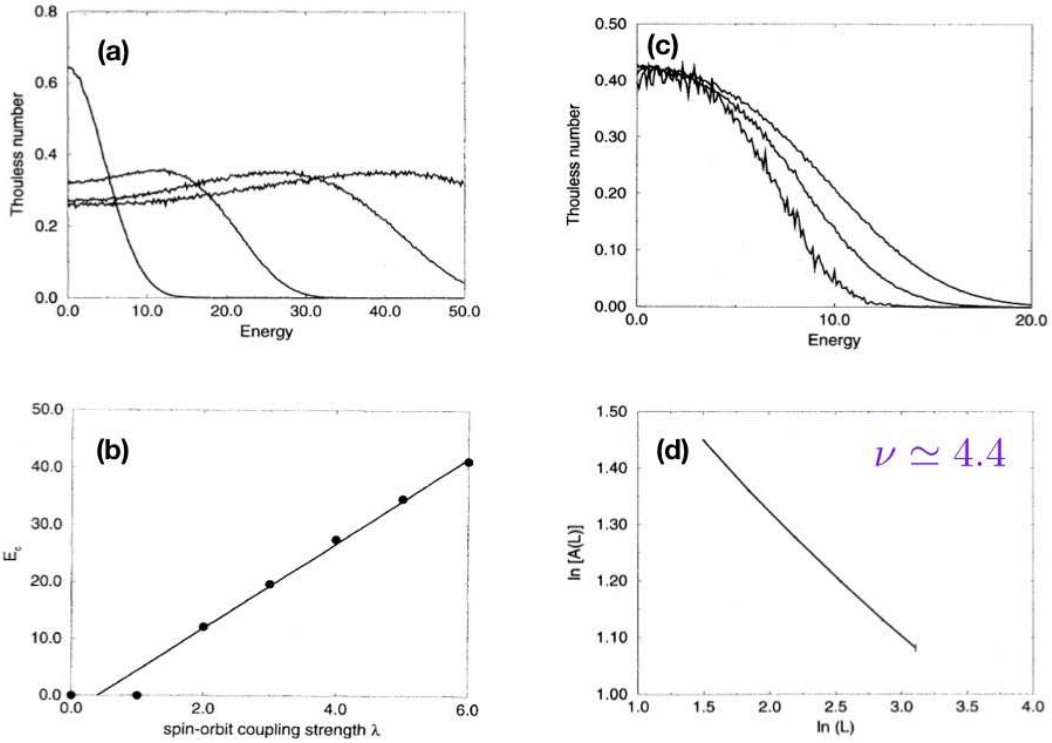


Figure 2.23: (a) Thouless number data for white noise SO scattering ($\zeta = 0$) for $N_\phi = 40, 60, 160,$ and 500. (b) Log-log plot of the integrated area under the Thouless number curves $T_L(E)$ *versus* system size. The resulting slope yields $\nu \approx 4.4$. (c) Thouless number for smooth SO scattering ($\zeta = 2$) as a function of energy for coupling strengths $\lambda = 0$, 2, 4, and 6. (d) Tracking of the Thouless number peak, with $E_c(\lambda) \propto \lambda W$. From C. B. Hanna *et al. op. cit.*.

Hanna *et al.*[61] (HAMG) studied the model

$$H = \Pi_0 \left[V(\boldsymbol{r}) + \lambda \, \boldsymbol{W}(\boldsymbol{r}) \cdot \boldsymbol{\sigma} \right] \Pi_0 \quad , \tag{2.116}$$

where Π_0 projects onto the $n = 0$ cyclotron Landau level, where $V(\boldsymbol{r})$ and $W_{x,y,z}(\boldsymbol{r})$ are Gaussian random fields of zero mean, satisfying

$$\begin{aligned} \big\langle V(\boldsymbol{r})\, V(\boldsymbol{r}') \big\rangle &= \frac{V^2}{2\pi\zeta^2} \, \exp\big(-|\boldsymbol{r} - \boldsymbol{r}'|^2/2\zeta^2 \big) \\ \big\langle W_\alpha(\boldsymbol{r})\, W_\beta(\boldsymbol{r}') \big\rangle &= \frac{W^2}{2\pi\zeta^2} \, \exp\big(-|\boldsymbol{r} - \boldsymbol{r}'|^2/2\zeta^2 \big) \, \delta_{\alpha\beta} \quad , \end{aligned} \tag{2.117}$$

where V and W are the respective strengths of the scalar and spin-dependent random potentials, and ζ is the correlation length of the disorder, assumed the same for both $V(\boldsymbol{r})$ and $\boldsymbol{W}(\boldsymbol{r})$. The limit $\zeta \to 0$ corresponds to Gaussian white noise, but is effectively smoothed on the scale of

[61] C. B. Hanna, D. P. Arovas, K. Mullen, and S. M. Girvin, *Phys. Rev. B* **52**, 5221 (1995).

the magnetic length ℓ due to the LLL projection Π_0. Note that there is no Zeeman term, corresponding to $g^* = 0$, hence the spins in the absence of the $\boldsymbol{W} \cdot \boldsymbol{\sigma}$ term are completely unresolved. In other words,

$$H_{j\alpha,j'\alpha'}(\theta_1,\theta_2) = \langle\, j \,|\, V(\boldsymbol{r}) \,|\, j' \,\rangle\, \delta_{\alpha\alpha'} + \lambda \,\langle\, j \,|\, \boldsymbol{W}(\boldsymbol{r}) \,|\, j' \,\rangle \cdot \boldsymbol{\sigma}_{\alpha\alpha'} \quad , \tag{2.118}$$

where (θ_1, θ_2) are the boundary condition Bloch phases. HAMG computed the *Thouless number*, which, for the $n^{\rm th}$ state in the spectrum, is defined to be

$$T_n = \overline{g(E_n)}\,\big|E_n(\pi,0) - E_n(0,0)\big| \quad , \tag{2.119}$$

where $\overline{g(E_n)}$ is the density of states at energy E_n, averaged over some width δE which contains many levels. The energy difference $\Delta E = E_n(\pi,0) - E_n(0,0)$ is the difference between values at antiperiodic and periodic boundary conditions in the θ_1 phase[62]. One finds $T_n = T(E_n)$ is a smooth function of the energy E_n after averaging over disorder. Assuming the scaling form

$$T_L(E) = f\big(\xi(E)/L\big) \equiv \tilde{f}(L^{1/\nu}|E|) \quad , \tag{2.120}$$

the area under the $T_L(E)$ curves behaves as

$$A(L) = \int\limits_{-\infty}^{\infty} dE\, T_L(E) = C\, L^{-1/\nu} \quad , \tag{2.121}$$

where $T_L(0)$ and C are independent of the system size L.

HAMG found that for smooth SO disorder scattering ($\zeta = 2$), there are indeed two fully resolved individual peaks of $T_L(E)$ located at $E = \pm E_{\rm c}$, with $E_{\rm c}$ independent of L for $L \gtrsim 40$ (see Fig. 2.22). The dependence of $E_{\rm c}$ on the SO coupling λ was found to be linear. Plotting $\ln A(L)$ *versus* $\ln L$, an exponent of $\nu \approx 2.5$ was extracted. Results white noise disorder ($\zeta = 0$) are shown in Fig. 2.23. The two peaks of $T_L(E)$ are poorly resolved, and an attempt to infer ν from the scaling *Ansatz* for $T_L(E)$ yields the approximately doubled value $\nu \simeq 4.4$. Much larger systems are apparently necessary in order to fully resolve the two peaks and to obtain the presumably correct value of $\nu \approx 2.5$.

2.3 Edge States

Recall the Hofstadter model from §1.6.2. Rather than applying doubly periodic boundary conditions in both x and y directions, consider the model on a cylinder with N_1 sites on each row parallel to the symmetry axis, and N_2 sites in the periodic direction. We will take $N_2 \to \infty$ but keep N_1 finite. A sketch is given in Fig. 2.24. The boundaries at the edges of the cylinder break

[62] One could equally well define the Thouless number with respect to boundary conditions in the θ_2 phase.

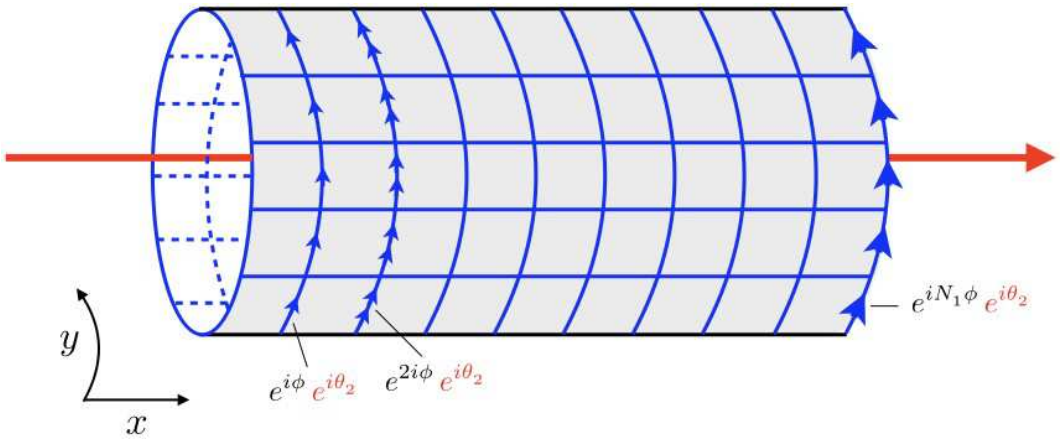


Figure 2.24: Hofstadter model on a cylinder, with flux ϕ per unit cell.

translational invariance in the x-direction, so the magnetic unit cell is now $N_x \times 1$ structural unit cells, and the Hamiltonian is

$$H(\theta_2, N_1, z) = -t \begin{pmatrix} 2\cos\theta_2 & 1 & 0 & \cdots & z^* \\ 1 & 2\cos(\theta_2+\phi) & 1 & & 0 \\ 0 & 1 & \ddots & & \vdots \\ \vdots & & & & 1 \\ z & 0 & \cdots & 1 & 2\cos(\theta_2 + N_1\phi) \end{pmatrix} , \tag{2.122}$$

which is an $N_1 \times N_1$ matrix. Here z controls the boundary condition in the x-direction, with $z = \exp(iN_1\theta_1/q)$ for periodic (toroidal) boundary conditions and $z = 0$ for open (cylindrical) boundary conditions.

For the infinite square lattice Hofstadter model with flux $\phi = 2\pi p/q$ per structural unit cell, TKNN showed that the Chern number C_r for the $r^{\rm th}$ subband is given by $C_r = t_r - t_{r-1}$, where $t_0 = 0$ and where t_r is determined by the solution of the Diophantine equation[63]

$$r = q\, s_r + p\, t_r \quad . \tag{2.123}$$

with $|t_r| \le \frac{1}{2}q$. For $p = 3$ and $q = 7$, one has $r = 7s_r + 3\,t_r$ and the solutions to the TKNN Diophantine equation are given by

$$\begin{aligned} &(s_1, t_1) = (1, -2) \quad , \quad (s_2, t_2) = (-1, 3) \quad , \quad (s_3, t_3) = (0, 1) \quad , \quad (s_4, t_4) = (1, -1) \\ &(s_5, t_5) = (2, -3) \quad , \quad (s_6, t_6) = (0, 2) \quad , \quad (s_7, t_7) = (1, 0) \quad . \end{aligned} \tag{2.124}$$

The t_r values are $(t_0, \ldots, t_7) = (0, -2, +3, +1, -1, -3, +2, 0)$ and thus the Chern numbers are found to be $(C_1, \ldots, C_7) = (-2, +5, -2, -2, -2, +5, -2)$.

[63] Our convention corresponds to TKNN's strong potential limit.

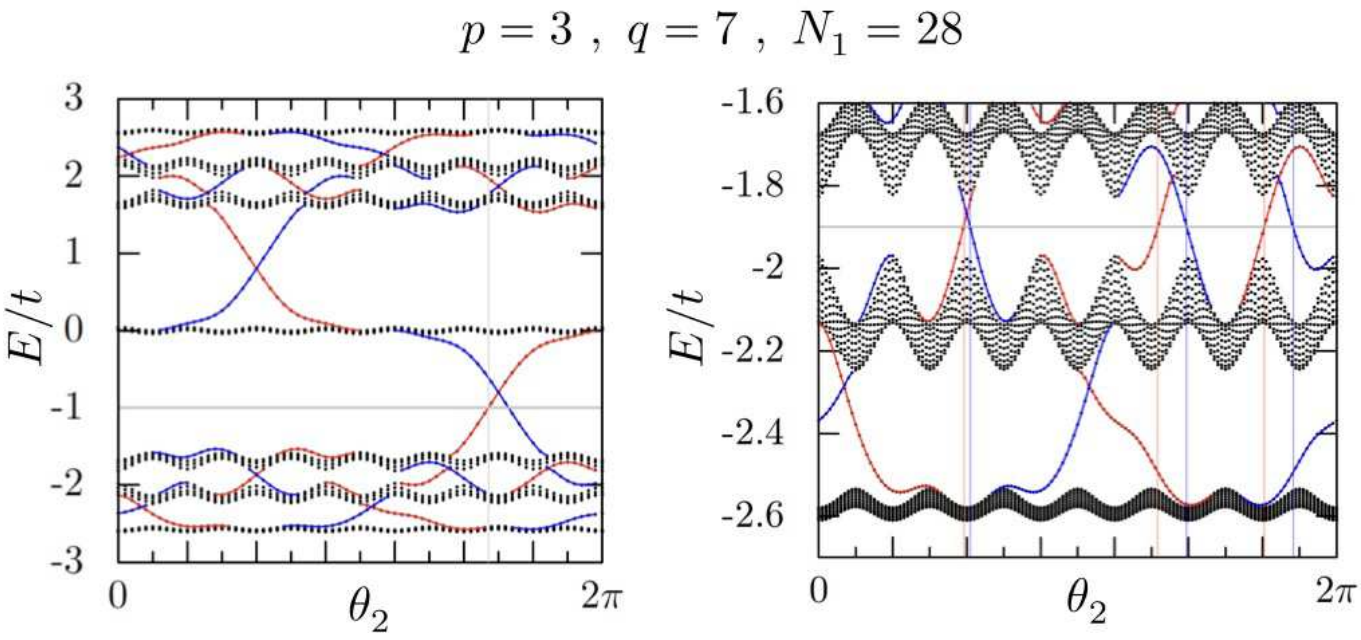


Figure 2.25: Left: bulk bands and edge states for the Hofstadter model with $p = 3$, $q = 7$, and $N_1 = 4q = 28$. Light vertical lines indicate θ_2 values corresponding to edge state crossings of the Fermi level (light horizontal line). Right: Detail of lowest three bulk bands and their associated edge states.

Consider now the results for the finite cylinder with $N_1 = 28$ shown in Fig. 2.25. Notice an essential difference relative to the bulk spectra: isolated energy levels traverse the gaps. These are *edge states* localized along either the left boundary of the cylinder (shown in red) or the right boundary (shown in blue). From the relation $\hbar v_2 = a\,\partial E/\partial\theta_2$, where a is the lattice constant, we see that the direction of each edge mode is determined by the sign of the slope of its dispersion curve whenever the Fermi level lies in a gap between bulk bands. Note that the modes associated with a given edge do not always propagate in the same direction as they do in the continuum (corresponding to the case $q \to \infty$), but can switch direction as the Fermi level is placed in consecutive bulk gaps.

Now look closely at Fig. 2.25 and note that there are two red edge levels propagating with $v_2 < 0$ when E_{F} lies in the gap between bands $r = 1$ and $r = 2$. This corresponds to the value $t_1 = -2$ obtained from the TKNN Diophantine equation above. When E_{F} lies in the gap between bulk bands $r = 2$ and $r = 3$, we see there are three red edge levels with $v_2 > 0$, corresponding to $t_2 = +3$, again in agreement with TKNN. Examining the next gap, we find $t_3 = +1$. In general, we observe the rule that *the Chern number C_r of the r^{th} band is the difference in the number of clockwise propagating states on the left edge when the Fermi level is changed from the $(r+1)^{\text{th}}$ to the r^{th} gap between bulk subbands*.

This correspondence between the bulk Chern number and the edge state structure is true in

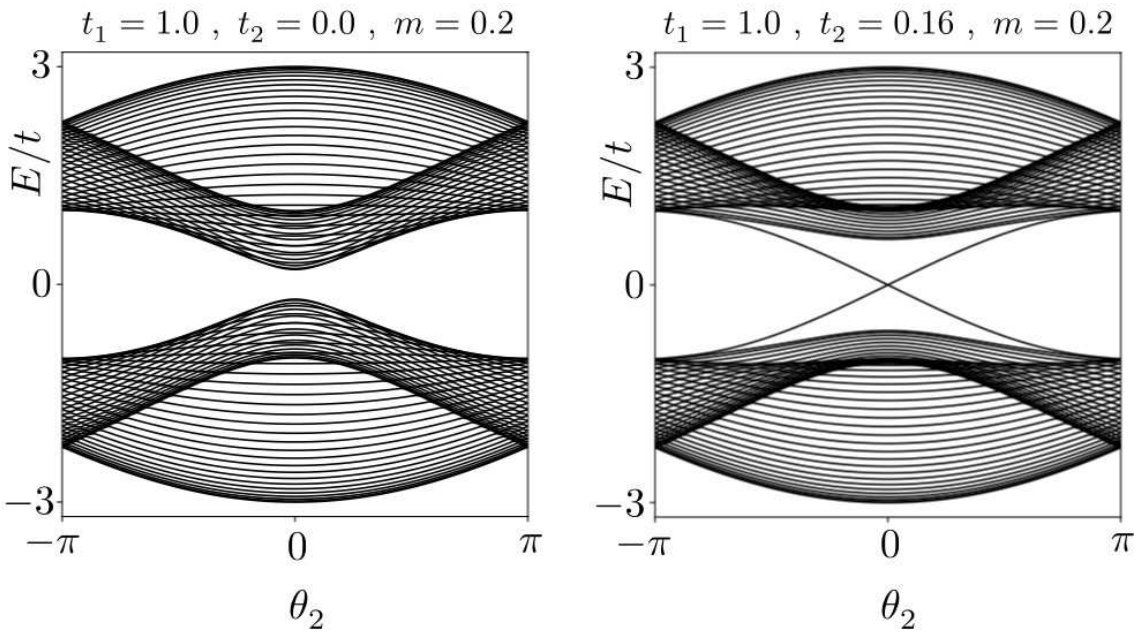


Figure 2.26: Bulk bands and edge states for the Haldane honeycomb lattice model with $t_1 = 1.0$ and $m = 0.2$. Left: $t_2 = 0$ (nontopological). Right: $t_2 = 0.16 > m/3\sqrt{3}$ (topological). Credit: https://topocondmat.org/w4_haldane/haldane_model.html.

general, and for another example consider the case of the Haldane honeycomb lattice model discussed in §1.7.3. When placed on a cylinder, the energy levels as a function of the Bloch phase θ_2 are depicted in Fig. 2.26, both in the nontopological $\big(|t_2| < |m|/3\sqrt{3}\big)$ and topological $\big(|t_2| > |m|/3\sqrt{3}\big)$ phases. Note how edge levels interpolating between the bulk bands are present in the topological phase, where the bulk band Chern numbers are $C_\pm = \mp 1$. Any lattice model with nonzero total Chern index when the Fermi level lies in a bulk gap is known as a *Chern insulator*.

2.3.1 Hatsugai's formulation

Yasuhiro Hatsugai in 1993 provided a particularly lucid description of the mathematics of edge states in lattice Chern insulators[64]. Consider a square lattice Chern insulator on a cylinder of dimensions $N_x \times N_y$, where y is the periodic direction, and where we take $N_y \to \infty$. Translational invariance in y guarantees that k_y is a good quantum number, and as usual we define $\theta_2 \equiv k_y a$, where a is the lattice constant. Let us fix our interest on the Hofstadter model with flux $\phi = 2\pi p/q$ per structural unit cell, and let $N_x = Jq$ where J is a positive integer. The

[64] Y. Hatsugai, *Phys. Rev. Lett.* **71**, 3697 (1993); Y. Hatsugai, *Phys. Rev. B* **48**, 11581 (1993).

wavefunction is described by the set of functions $\{\psi_n(\theta_2)\}$, where $n \in \{0, \ldots, N_x\}$.

$$\begin{aligned} \text{toroidal} \ : \ \psi_{n+N_x}(\theta_2) &= \psi_n(\theta_2) \ \forall \ n \\ \text{cylindrical} \ : \ \psi_0(\theta_2) &= \psi_{N_x}(\theta_2) = 0 \quad . \end{aligned} \tag{2.125}$$

In the toroidal case, choosing a $q \times 1$ magnetic unit cell, we have $\psi_{n+q}(\theta_2) = e^{i\theta_1}\psi_n(\theta_2)$ with $\theta_1 = 2\pi j/J$ and $j \in \{1, \ldots, J\}$.

The lattice Schrödinger equation for the Hofstadter model is

$$-t\,\psi_{n-1} - 2t\cos(n\phi + \theta_2)\,\psi_n - t\,\psi_{n+1} = E\,\psi_n \quad , \tag{2.126}$$

which may be restated as

$$\begin{pmatrix} \psi_{n+1} \\ \psi_n \end{pmatrix} = \overbrace{\begin{pmatrix} -\varepsilon - 2\,\cos(n\phi + \theta_2) & -1 \\ 1 & 0 \end{pmatrix}}^{R_n(\varepsilon, \theta_2, \phi)} \begin{pmatrix} \psi_n \\ \psi_{n-1} \end{pmatrix} \quad , \tag{2.127}$$

with $\varepsilon \equiv E/t$. We define the transfer matrix

$$M(\varepsilon) \equiv R_q(\varepsilon) R_{q-1}(\varepsilon) \cdots R_1(\varepsilon) \quad , \tag{2.128}$$

where we suppress notation of θ_2 and ϕ for clarity. The full transfer matrix across the cylinder is then $M^J(\varepsilon)$, and given our boundary condition $\psi_0(\theta_2) = 0$, we have

$$\begin{pmatrix} \psi_{Jq+1} \\ \psi_{Jq} \end{pmatrix} = M^J(\varepsilon) \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad . \tag{2.129}$$

This requires $\big[M^J(\varepsilon)\big]_{21} = 0$, which is a degree $N_x - 1 = Jq - 1$ polynomial equation in ε for each θ_2. Writing

$$N_x - 1 = (N_x - q) + (q - 1) = (J - 1)\,q + (q - 1) \quad , \tag{2.130}$$

we have that these $N_x - 1$ energy eigenstates for each θ_2 are grouped into q bands each with $(J-1)$ states, plus $(q-1)$ mid-gap states, which are the edge states. The condition $\big[M^J(\varepsilon)\big]_{21} = 0$ says that $M^J(\varepsilon)$ is a 2×2 upper-triangular matrix. It is satisfied by the simpler condition $M_{21}(\varepsilon) = 0$, which is an order $q - 1$ polynomial equation in ε, since any product of upper-triangular matrices is upper-triangular. It turns out that this condition sets the values of the $q - 1$ edge state energies, $\varepsilon = \mu_l$ with $l \in \{1, \ldots, q - 1\}$. The r^{th} edge state energy μ_j lies in the gap between bulk bands r and $r + 1$. Furthermore, since $\psi_0 = 0$ and we may choose $\psi_1 = 1$ (this is scaled by whatever normalization we may apply) for each r, we have

$$\psi^{(r)}_{kq+1}(\mu_r) = \big[M_{11}(\mu_r)\big]^k \tag{2.131}$$

and therefore we conclude

$$\begin{aligned} &\bullet \ \big|M_{11}(\mu_r)\big| < 1 \quad \Rightarrow \quad \psi^{(r)}_i \text{ localized on left edge } (i \approx 1) \\ &\bullet \ \big|M_{11}(\mu_r)\big| > 1 \quad \Rightarrow \quad \psi^{(r)}_i \text{ localized on right edge } (i \approx N_x - 1) \quad . \end{aligned} \tag{2.132}$$

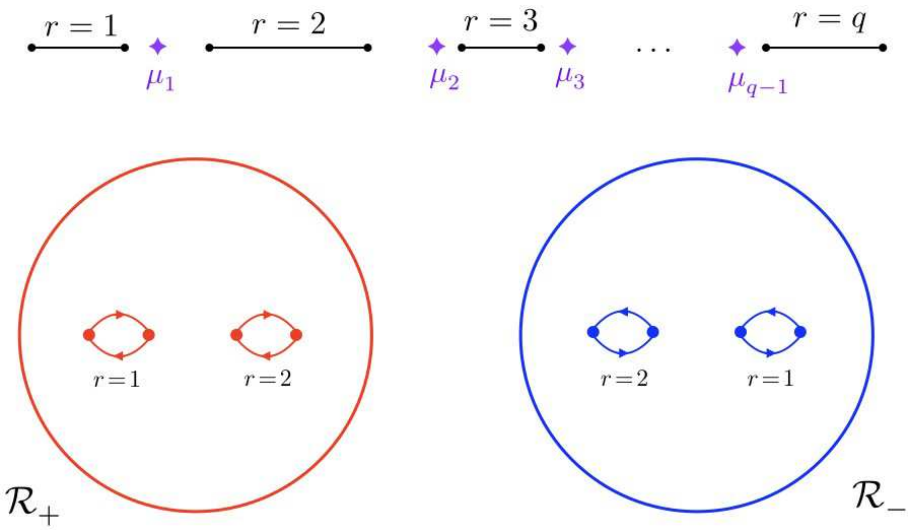


Figure 2.27: Hatsugai's construction of the genus $g = q - 1$ Riemann surface.

When $\left|M_{11}(\mu_r)\right| = 1$, the edge level merges with the bulk and there is no exponential localization.

With doubly periodic (*i.e.* toroidal) boundary conditions, the Bloch condition is

$$\begin{pmatrix}\psi_{q+1}\\ \psi_q\end{pmatrix} = M(\varepsilon)\begin{pmatrix}\psi_1\\ \psi_0\end{pmatrix} = e^{i\theta_1}\begin{pmatrix}\psi_1\\ \psi_0\end{pmatrix} \quad . \tag{2.133}$$

Following Hatsugai, we now analytically continue $\varepsilon \to z \in \mathbb{C}$ and we define the phase $\rho \equiv \exp(i\theta_1)$. Solving for $\rho(z)$, we have $\det\big(\rho - M(z)\big) = \rho^2 - T(z)\,\rho + 1 = 0$, where $T(z) = \mathrm{Tr}\, M(z)$. Note that $\det M(z) = 1$ since $\det R_n(z) = 1$ for all n. The solution is

$$\rho(z) = \tfrac{1}{2}T(z) \pm \tfrac{1}{2}\sqrt{T^2(z) - 4} \quad . \tag{2.134}$$

Furthermore, we have

$$\psi_0 = -\frac{M_{21}\,\psi_1}{M_{22} - \rho} \qquad \Rightarrow \qquad \psi_q = -\frac{\rho M_{21}}{M_{22} - \rho} = \frac{1}{M_{12}}\,\rho\,(\rho - M_{11}) \quad . \tag{2.135}$$

Define $\omega(z) = \sqrt{T^2(z) - 4}$. The branch cuts in $\omega(z)$ define the bulk energy bands, where $T^2(z) < 4$ and $z = \varepsilon \in \mathbb{R}$, which entails $|\rho(\varepsilon)| = 1$. Therefore, we can write

$$\omega(z) = \sqrt{(z - \lambda_1)(z - \lambda_2)\cdots(z - \lambda_q)} \tag{2.136}$$

where the $r^{\rm th}$ bulk band energies satisfy $\varepsilon \in [\lambda_{2r-1}, \lambda_{2r}]$. We may compactify, taking all values of z for $|z| \to \infty$ to a single point. The need to specify a sign for $\omega(z)$, which is a square root, means we have two spheres, $\mathcal{R}_+$ and $\mathcal{R}_-$, each with q branch cuts corresponding to the bulk

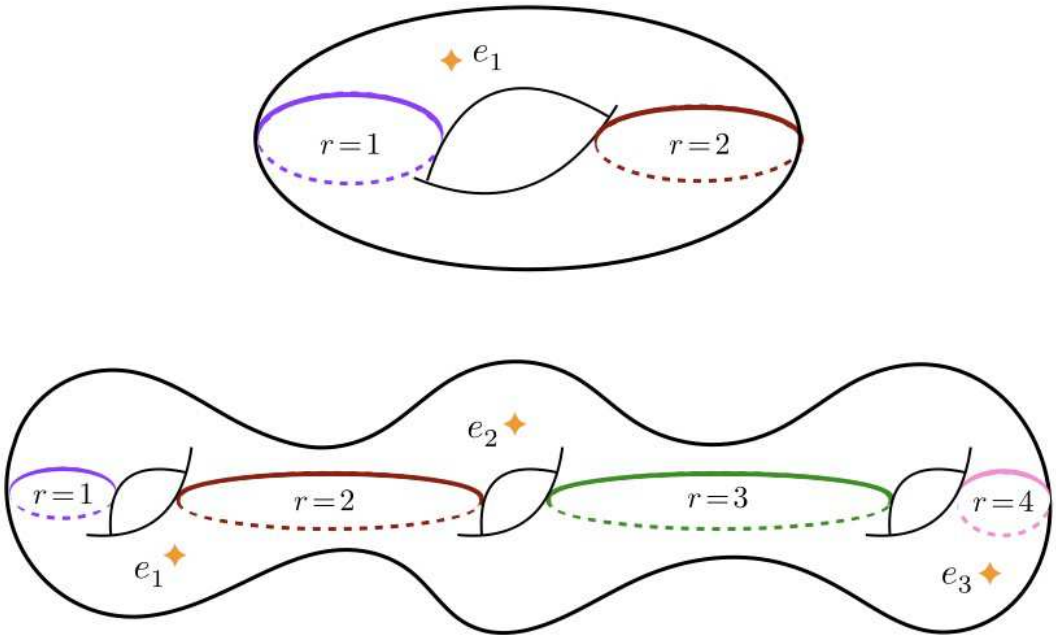


Figure 2.28: Genus $g = 1$ and $g = 3$ Riemann surface with q bulk bands and $q - 1$ edge states for each value of θ_2.

bands. The branch of $\omega(z)$ is defined such that $\omega(z \to -\infty) \gtrless 0$ on $\mathcal{R}_\pm$. The spheres $\mathcal{R}_\pm$ are glued together as shown in Fig. 2.27 to create a genus $g = q - 1$ Riemann surface, $\Sigma_g(\theta_2)$ for each θ_2, which is the complex energy surface for the Hofstadter model (see Fig. 2.28). Note that g is the number of gaps, and hence the number of edge modes.

Each edge state must satisfy $\psi_q^{(r)}(\mu_r) = 0$. There are g such solutions on $\Sigma_g(\theta_2)$. As θ_2 is varied, $\mu_r(\theta_2)$ moves around the $r^{\rm th}$ hole in an oriented loop. The winding number of this loop, I_r, determines the Hall conductance, with the Chern number of the $r^{\rm th}$ band being given by

$$C_r = I_{r-1} - I_r \qquad \Rightarrow \qquad \sigma_{xy} = \frac{e^2}{h} \sum_{j=1}^{r} C_j = -\frac{e^2}{h} I_r \quad , \tag{2.137}$$

where r is the label of the highest-lying filled band.

2.3.2 Qi-Wu-Zhang picture

Recall that the raisin bagel, while a culinary abomination and an outrageous form of cultural appropriation of my people, nevertheless furnishes a useful paradigm for the Wigner - von Neumann theorem, which says that accidental degeneracy for complex Hamiltonians has co-

dimension three[65]. The raisin bagel corresponds to a three-dimensional *filled torus*, parameterized by the two Bloch phases (θ_1, θ_2) and a third radial coordinate $r \in [0, 1]$. Degeneracies of two neighboring bands, $E_n(\theta_1, \theta_2, r)$ and $E_{n+1}(\theta_1, \theta_2, r)$, occur at discrete points $(\theta_1^*, \theta_2^*, r^*)$ within the bagel. We identify these points as the raisins.

In the picture of Qi, Wu, and Zhang[66] (QWZ), the radial coordinate r is a multiplicative factor in the hopping amplitudes on all links coupling sites with x-values N_x and 1. The Hamiltonian is taken to be

$$H = -\sum_{\langle \boldsymbol{R}\boldsymbol{R}' \rangle} \sum_{\alpha,\beta} \left(t_{\boldsymbol{R}\boldsymbol{R}',\alpha\beta}\, c^\dagger_{\boldsymbol{R}\alpha}\, c_{\boldsymbol{R}'\beta} + \text{H.c.} \right) + V \quad , \tag{2.138}$$

where V conserves local particle number $n_{\boldsymbol{R}} = \sum_\alpha c^\dagger_{\boldsymbol{R}\alpha}\, c_{\boldsymbol{R}\alpha}$ at every site $\boldsymbol{R}$, and may describe disorder or locally n-preserving interactions. On the links which straddle the horizontal and vertical "edges" Σ_{h} and Σ_{v} of the toroidal base space, we take

$$\begin{aligned} t_{\boldsymbol{R}\boldsymbol{R}'} &\longrightarrow r\, t_{\boldsymbol{R}\boldsymbol{R}'}\, \exp(i\theta_1 Q_1) \quad \text{horizontal edge} \\ t_{\boldsymbol{R}\boldsymbol{R}'} &\longrightarrow t_{\boldsymbol{R}\boldsymbol{R}'}\, \exp(i\theta_2 Q_2) \quad \text{vertical edge} \quad , \end{aligned} \tag{2.139}$$

where $Q_{1,2}$ are Hermitian matrices satisfying $\exp(2\pi i Q_j) = 1$, and where $t_{\boldsymbol{R}\boldsymbol{R}'}$ is for each link $\langle \boldsymbol{R}\boldsymbol{R}' \rangle$ a matrix with internal indices α and β, *viz.* $t_{\boldsymbol{R}\boldsymbol{R}',\alpha\beta}$ in Eqn. 2.138. Thus $H = H(\theta_1, \theta_2, r)$ has a three-dimensional parameter space, the filled torus, which interpolates between a cylinder at $r = 0$ and a torus at $r = 1$.

The Berry gauge connection for the n^{th} energy band is

$$A^{(n)}_\mu(\theta_1, \theta_2, r) = -i \left\langle\, \psi_n(\theta_1, \theta_2, r) \,\middle|\, \frac{\partial}{\partial \theta^\mu} \,\middle|\, \psi_n(\theta_1, \theta_2, r) \,\right\rangle \tag{2.140}$$

and the associated Berry curvature is

$$\Omega^{(n)}_{\mu\nu}(\theta_1, \theta_2, r) = \frac{\partial A^{(n)}_\nu}{\partial \theta^\mu} - \frac{\partial A^{(n)}_\mu}{\partial \theta^\nu} \quad . \tag{2.141}$$

At fixed r, integrating around a contour $\mathcal{C}$ on the (θ_1, θ_2) torus, one has

$$\Phi_n(\mathcal{C}, r) = \oint_{\mathcal{C}} d\theta^\mu\, A^{(n)}_\mu(\theta_1, \theta_2, r) \quad . \tag{2.142}$$

For notational clarity, we henceforth suppress the band index n.

[65] Seeded bagels are perfectly acceptable and indeed delicious. But bagels should not be defiled with cinnamon or sugar or fruit. That's the sort of thing that Hitler would probably do.

[66] X.-L. Qi, Y.-S. Wu, and S.-C. Zhang, *Phys. Rev. B* **74**, 045125 (2006).

In general, the connection cannot be defined globally, and instead only on patches. To this end, we define, for $r = 1$, $A_1^{\rm I}(\theta_1, \theta_2, r = 1) = A_1^{\rm II}(\theta_1, \theta_2, r = 1) = 0$, and

$$A_2^{\rm I}(\theta_1, \theta_2, , r = 1) = \int\limits_0^{\theta_1} d\theta_1' \; \Omega_{12}(\theta_1', \theta_2) \quad \text{for } \theta_1 \in (0, 2\pi)$$
$$A_2^{\rm II}(\theta_1, \theta_2, , r = 1) = \int\limits_{-\pi}^{\theta_1} d\theta_1' \; \Omega_{12}(\theta_1', \theta_2) \quad \text{for } \theta_1 \in (-\pi, \pi) \quad . \tag{2.143}$$

Note the discontinuity in $A_2(\theta_1, \theta_2, , r = 1)$ at $\theta_1 = 0$ and $\theta_1 = \pi$. Thus, the (θ_1, θ_2) torus $\mathbb{T}^2$ is covered by two cylinders $\theta_1 \neq 0$ and $\theta_1 \neq \pi$. The Chern number is given by

$$C = \lim_{\epsilon \to 0} \frac{1}{2\pi} \int\limits_{\epsilon}^{2\pi - \epsilon} d\theta_1 \; \frac{\partial}{\partial \theta_1} \left(\int\limits_0^{2\pi} d\theta_2 \; A_2^{\rm I}(\theta_1, \theta_2, , r = 1) \right) \quad . \tag{2.144}$$

Following QWZ, we may define the quantity

$$\chi(\theta_1, r) \equiv \int\limits_0^{2\pi} d\theta_2 \; A_2^{\rm I}(\theta_1, \theta_2, r) \tag{2.145}$$

and the phase $\Upsilon(\theta_1, r) \equiv \exp\big(i\chi(\theta_1, r)\big)$, so that

$$C = \frac{1}{2\pi} \oint\limits_{r=1} d\Upsilon \; \Upsilon^{-1} = \frac{1}{2\pi} \int\limits_{0^+}^{2\pi^-} d\theta_1 \; \frac{\partial \chi(\theta_1, r)}{\partial \theta_1} \quad . \tag{2.146}$$

Although $\chi(\theta_1, r)$ jumps by $2\pi C$ across $\theta_1 \in [0, 2\pi]$, the function $\Upsilon(\theta_1, r)$ is everywhere single-valued and well-behaved. Since $H = H(r\, e^{i\theta_1 Q_1}\, ,\, e^{i\theta_2 Q_2})$, when $r = 0$ at fixed θ_2, the Hamiltonian is the same for all θ_1. Thus, (r, θ_1) may be viewed as 2D polar coordinates, with $r = 0$ the origin. It follows that if $C \neq 0$, there must be a vortex singularity somewhere within the unit disk $\{(r, \theta_1) \,|\, 0 \leq r \leq 1 \,,\, 0 \leq \theta \leq 2\pi\}$. Since $\chi(\theta_1, r)$ is well-defined provided the state $|\, \psi_n \,\rangle$ is nondegenerate, we conclude that one or both of the gaps $E_n - E_{n-1}$ or $E_{n+1} - E_n$ must collapse at some point in the interior of the disk if $C_n \neq 0$. Thus,

- ⋆ *Whenever $C_n \neq 0$ for the $r = 1$ system, there must exist one or more points $(\theta_1^*, \theta_2^*, r^*)$ with $r^* < 1$ where the state $|\, \psi_n(\theta_1^*, \theta_2^*, r^*) \,\rangle$ is degenerate with one of $|\, \psi_{n\pm1}(\theta_1^*, \theta_2^*, r^*) \,\rangle$.*

This is essentially a restatement of Wigner - von Neumann. Note that $r < 1$ corresponds to weakened inter-edge tunneling.

Chapter 3

Fractional Quantum Hall Effect

3.1 Many-Body States in the Lowest Landau Level

3.1.1 Introduction

Transport experiments in the quantum Hall regime reveal QH plateaus at fractional values of $\sigma_{xy} = \nu\, e^2/h$ with $\nu = p/q$ a rational fraction, principally with q odd. This corresponds to fractional filling of a Landau level. Recall that electrons are fermions, and a many-body fermionic wavefunction must be totally antisymmetric with respect to change of labels, *viz.*

$$\Psi(\boldsymbol{\xi}_{\sigma(1)}\,,\,\ldots\,,\,\boldsymbol{\xi}_{\sigma(N)}) = \mathsf{sgn}(\sigma)\,\Psi(\boldsymbol{\xi}_1\,,\,\ldots\,,\,\boldsymbol{\xi}_N) \quad , \tag{3.1}$$

where $\boldsymbol{\xi}_j = (\boldsymbol{r}_j, \tau_j)$ is a compound variable including the spatial coordinates and spin polarization ($\tau_j = \pm 1$) for the $j^{\rm th}$ electron, and where $\mathsf{sgn}(\sigma) \equiv (-1)^\sigma$ is the sign of the permutation $\sigma \in \mathcal{S}_N$. Initially we will presume that the Zeeman field polarizes all the electrons into the same spin state with $\tau_j = +1$ for all j. In this case we only need concern ourselves with the spatial coordinates $\{\boldsymbol{r}_j\}$.

One way to construct such a totally antisymmetric state is via the Slater determinant,

$$\Psi(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_N) = \det\big\{\varphi_i(\boldsymbol{r}_j)\big\} = \det\begin{pmatrix} \varphi_1(\boldsymbol{r}_1) & \varphi_1(\boldsymbol{r}_2) & \cdots & \varphi_1(\boldsymbol{r}_N) \\ \varphi_2(\boldsymbol{r}_1) & \varphi_2(\boldsymbol{r}_2) & \cdots & \varphi_2(\boldsymbol{r}_N) \\ \vdots & & \ddots & \vdots \\ \varphi_N(\boldsymbol{r}_1) & \cdots & \cdots & \varphi_N(\boldsymbol{r}_N) \end{pmatrix} \quad , \tag{3.2}$$

Here $\big\{\varphi_i(\boldsymbol{r})\big\}$ is a basis of single particle wavefunctions. Recall that in the LLL, in the symmetric gauge $\boldsymbol{A} = \frac{1}{2}B(y, -x)$, for which $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$, all the wavefunctions are of the restricted form $\psi(\boldsymbol{r}) = f(z)\,\exp(-z\bar{z}/4\ell^2)$, where $f(z)$ is an analytic function in $z = x + iy$, meaning $\bar{\partial} f(z) = 0$,

where $\bar\partial \equiv \partial_{\bar z} = \frac{1}{2}(\partial_x + i\partial_y)$. In the LLL, then, the most general N-electron wavefunction is of the form

$$\Psi(\boldsymbol{r}_1\,,\,\ldots\,,\,\boldsymbol{r}_N) = F(z_1,\ldots,z_N)\prod_{j=1}^{N}\exp\bigl(-z_j\bar z_j/4\ell^2\bigr) \quad , \tag{3.3}$$

where F is analytic in all its arguments. Now recall the angular momentum basis for the LLL,

$$\varphi_m(\boldsymbol{r}) = \frac{1}{\sqrt{2\pi\ell^2 m!}}\left(\frac{z}{\sqrt{2}\,\ell}\right)^m e^{-z\bar z/4\ell^2} \quad . \tag{3.4}$$

Without normalization, we have the analytic factor $f_m(z) = z^m$, and forming a Slater determinant among N electrons in the angular momentum states $m \in \{0, 1, \ldots, N-1\}$, we have

$$F(z_1,\ldots,z_N) = \mathsf{det}\bigl\{z_j^m\bigr\} = \mathsf{det}\begin{pmatrix} z_1^0 & z_2^0 & \cdots & z_N^0 \\ z_1^1 & z_2^1 & \cdots & z_N^1 \\ \vdots & & \ddots & \vdots \\ z_1^{N-1} & \cdots & \cdots & z_N^{N-1} \end{pmatrix} \quad . \tag{3.5}$$

Clearly $F(Z)$ is a homogeneous polynomial in its arguments $Z = \{z_1,\ldots,z_N\}$, which says $F(\lambda Z) = \lambda^{\mathsf{deg}\,F}F(Z)$. Since the $k^{\rm th}$ row of $F(\lambda Z)$ is multiplied by z^{k-1}, we have

$$\mathsf{deg}\,F = \sum_{k=1}^{N}(k-1) = \tfrac{1}{2}N(N-1) \quad . \tag{3.6}$$

Furthermore, since $F(Z)$ is totally antisymmetric, it must vanish whenever $z_i = z_j$ for all $i \neq j$. Thus, the product

$$V(Z) \equiv \prod_{i>j}(z_i - z_j) \tag{3.7}$$

must be a factor of $F(Z)$. But since there are $\frac{1}{2}N(N-1)$ terms in the product for $V(Z)$, we must have that $F(Z) = CV(Z)$, where C is a constant. Since the coefficient of the term $z_1^0 z_2^1\cdots z_N^{N-1}$ in both $F(Z)$ and $V(Z)$ is 1, we conclude $C = 1$ and hence $F(Z) = V(Z)$, which is called the *Vandermonde determinant*. It corresponds to the holomorphic part of the N-body LLL wavefunction where each of the lowest N angular momentum states, *i.e.* with $m \in \{0,\ldots,N-1\}$, is filled, with no holes. *The Vandermonde determinant holomorphic factor corresponds to a filled Landau level.* The many-body normalization integral is

$$\int d^2r_1\cdots\int d^2r_N\,\bigl|V(z_1,\ldots,z_N)\bigr|^2\exp\left(-\frac{1}{2\ell^2}\sum_{i=1}^{N}|z_i|^2\right) = N!\prod_{m=0}^{N-1}\Bigl[2\pi\ell^2\bigl(\sqrt{2}\,\ell\bigr)^m m!\Bigr] \quad . \tag{3.8}$$

3.1.2 Second quantization

With an orthonormal set of single particle wavefunctions $\{\varphi_\alpha(\boldsymbol{r}_i)\}$, the normalized Slater determinant state is given by

$$\Psi_{\alpha_1\cdots\alpha_N}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) = \frac{1}{\sqrt{N!}}\sum_{\sigma\in\mathcal{S}_N}\mathsf{sgn}(\sigma)\,\varphi_{\alpha_{\sigma(1)}}(\boldsymbol{r}_1)\cdots\varphi_{\alpha_{\sigma(N)}}(\boldsymbol{r}_N) \quad . \tag{3.9}$$

We define the state

$$|\,\alpha_1,\ldots,\alpha_N\,\rangle = \frac{1}{\sqrt{N!}}\sum_{\sigma\in\mathcal{S}_N}\mathsf{sgn}(\sigma)\,|\,\alpha_{\sigma(1)}\,\rangle\otimes\cdots\otimes|\,\alpha_{\sigma(N)}\,\rangle \equiv c^\dagger_{\alpha_N}\cdots c^\dagger_{\alpha_1}\,|\,0\,\rangle \tag{3.10}$$

in which case $\Psi_{\alpha_1\cdots\alpha_N}(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) = \langle\,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N\,|\,\alpha_1,\ldots,\alpha_N\,\rangle$. Here $\{c_\alpha\,,\,c^\dagger_\beta\} = \delta_{\alpha\beta}$ are the canonical anticommutation relations for fermionic annihilation (c_α) and creation ($c^\dagger_\beta$) operators.

The second quantized Hamiltonian is written as $\hat{H} = \hat{T} + \hat{U} + \hat{V}$. The kinetic energy is

$$\hat{T} = \sum_{\alpha,\beta}\langle\,\alpha\,|\,t\,|\,\beta\,\rangle\,c^\dagger_\alpha\,c_\beta \quad , \tag{3.11}$$

where

$$\langle\,\alpha\,|\,t\,|\,\beta\,\rangle = \int d^dr\;\varphi^*_\alpha(\boldsymbol{r})\,t\big(\boldsymbol{r},\boldsymbol{\nabla}\big)\,\varphi_\beta(\boldsymbol{r}) \quad , \tag{3.12}$$

where $t(\boldsymbol{r},\boldsymbol{\nabla})$ is the single particle kinetic energy operator, and is often a function of the vector derivative $\boldsymbol{\nabla}$ alone, is $t = -\frac{\hbar^2}{2m}\boldsymbol{\nabla}^2$. Of course for a particle in a magnetic field, we have that $t = \frac{\hbar^2}{2m}\big(-i\boldsymbol{\nabla} + \frac{e}{\hbar c}\boldsymbol{A}\big)^2$. A single particle potential $u(\boldsymbol{r})$ gives rise to the second quantized contribution

$$\hat{U} = \sum_{\alpha,\beta}\langle\,\alpha\,|\,u\,|\,\beta\,\rangle\,c^\dagger_\alpha\,c_\beta \quad , \tag{3.13}$$

where

$$\langle\,\alpha\,|\,u\,|\,\beta\,\rangle = \int d^dr\;\varphi^*_\alpha(\boldsymbol{r})\,u(\boldsymbol{r})\,\varphi_\beta(\boldsymbol{r}) \quad , \tag{3.14}$$

Finally, the two-body potential is given in second quantized form as

$$\hat{V} = \frac{1}{2}\sum_{\alpha,\beta,\gamma,\delta}\langle\,\alpha\beta\,|\,v\,|\,\gamma\delta\,\rangle\;c^\dagger_\alpha\,c^\dagger_\beta\,c_\delta\,c_\gamma \quad , \tag{3.15}$$

where

$$\langle\,\alpha\beta\,|\,v\,|\,\gamma\delta\,\rangle = \int d^dr_1\int d^dr_2\;\varphi^*(\boldsymbol{r}_1)\,\varphi^*(\boldsymbol{r}_2)\,v(\boldsymbol{r}_1-\boldsymbol{r}_2)\,\varphi_\delta(\boldsymbol{r}_2)\,\varphi_\gamma(\boldsymbol{r}_1) \quad . \tag{3.16}$$

The field operator is given by

$$\psi(\boldsymbol{r}) = \sum_\alpha\varphi_\alpha(\boldsymbol{r})\,c_\alpha \qquad , \qquad \psi^\dagger(\boldsymbol{r}) = \sum_\alpha\varphi^*_\alpha(\boldsymbol{r})\,c^\dagger_\alpha \quad . \tag{3.17}$$

Thus,

$$\big\{\psi(\boldsymbol{r})\,,\,\psi^\dagger(\boldsymbol{r}')\big\} = \sum_\alpha \varphi_\alpha^*(\boldsymbol{r})\,\varphi_\alpha(\boldsymbol{r}') = \delta(\boldsymbol{r}-\boldsymbol{r}') \qquad \text{(entire Hilbert space)}$$
$$= \frac{1}{2\pi\ell^2}\,e^{i\,\mathsf{Im}\,(\bar z z')/2\ell^2}\,e^{-|z-z'|^2/4\ell^2} \quad \text{(LLL only)} \quad . \tag{3.18}$$

As an example of the second quantized formalism, consider the density operator

$$n(\boldsymbol{r}) = \psi^\dagger(\boldsymbol{r})\,\psi(\boldsymbol{r}) = \sum_{m_1}\sum_{m_2}\varphi_{m_1}^*(\boldsymbol{r})\,\varphi_{m_2}(\boldsymbol{r})\,c_{m_1}^\dagger c_{m_2} \quad . \tag{3.19}$$

where we use the angular momentum basis. Let $|\,\Psi_1\,\rangle = \prod_{m=0}^{N_\phi-1} c_m^\dagger\,|\,0\,\rangle$ denote the filled Landau level, where N_ϕ is the Landau level degeneracy and $N = N_\phi$ is the number of electrons. Then

$$\langle\,\Psi_1\,|\,c_{m_1}^\dagger c_{m_2}\,|\,\Psi_1\,\rangle = \delta_{m_1,m_2} \quad , \tag{3.20}$$

and therefore

$$n(\boldsymbol{r}) = \langle\,\Psi_1\,|\,\psi^\dagger(\boldsymbol{r})\,\psi(\boldsymbol{r})\,|\,\Psi_1\,\rangle = \sum_{m=0}^{N_\phi-1}\big|\varphi_m(\boldsymbol{r})\big|^2 = \frac{1}{2\pi\ell^2}\sum_{m=0}^{N_\phi-1}\frac{1}{m!}\left(\frac{|z|^2}{2\ell^2}\right)^m e^{-|z|^2/2\ell^2} \quad . \tag{3.21}$$

In the limit $N_\phi\to\infty$, we have $n(\boldsymbol{r})\to 1/2\pi\ell^2$, the number density of a filled Landau level. For finite N_ϕ, the electron density is described by a droplet of radius R, where $\pi R^2 = 2\pi\ell^2 N_\phi$. To see this, let $\zeta\equiv|z|^2/2\ell^2$, so $\nu(\zeta)\equiv 2\pi\ell^2\,n(\boldsymbol{r}) = e^{-\zeta}\sum_{m=0}^M \zeta^m/m!$ where $M\equiv N_\phi-1$. Thus we have $d\nu/d\zeta = -e^{-\zeta}\,\zeta^M/M!$ which is maximized in magnitude at $\zeta=M$, where for large M it takes the value $-1/\sqrt{2\pi M}$. Now using the chain rule we obtain $(d\nu/dr)_{\rm min} = -1/2\sqrt{\pi}\ell$. Thus, $\nu(r)$ drops from $\nu\approx 1$ inside the droplet, *i.e.* $r<R=(2N_\phi)^{1/2}\,\ell$, to $\nu\approx 0$ outside the droplet on a distance scale $\Delta r\sim\ell$.

We can carry out the same computation in the Landau basis, where in the $n=0$ LL

$$\psi_k(x) = \big(\sqrt{\pi}\,\ell L\big)^{-1/2}\,e^{iky}\,e^{-(x-k\ell^2)^2/2\ell^2} \quad . \tag{3.22}$$

Suppose we fill all states with $k<0$, so

$$n(x) = L\int_{-\infty}^{0}\frac{dk}{2\pi}\,\big|\psi_k(x)\big|^2 = \frac{1}{4\pi\ell^2}\,\mathsf{erfc}(x/\ell) \quad , \tag{3.23}$$

where

$$\mathsf{erfc}(z) = \frac{2}{\sqrt{\pi}}\int_z^\infty dt\,e^{-t^2} = 1-\mathsf{erf}(z) \quad . \tag{3.24}$$

Note $\mathsf{erfc}(-\infty) = 1$ while $\mathsf{erfc}(0) = \frac{1}{2}$ and $\mathsf{erfc}(\infty) = 0$. Thus there is an edge at $x = 0$, across which the electron density drops from $n(-\infty) = 1/2\pi\ell^2$ to $n(\infty) = 0$ within an interval $\Delta x \sim \ell$. If we further assume a neutralizing background of number density $\Theta(-x)/2\pi\ell^2$, then the total charge density in units of the electron charge is given by

$$\rho(x) \equiv n(x) - \frac{\Theta(-x)}{2\pi\ell^2} = \frac{1}{4\pi\ell^2}\,\mathsf{erfc}(x/\ell)\,\mathsf{sgn}(x) \quad . \tag{3.25}$$

Thus there is overall charge neutrality, *i.e.* $\int\limits_{-\infty}^{\infty} dx\,\rho(x) = 0$, and we may define a dipole moment per unit length

$$\delta = \int\limits_{-\infty}^{\infty} dx\,x\,\rho(x) = \frac{1}{8\pi} \quad . \tag{3.26}$$

For our next trick, let's evaluate the expression

$$\begin{aligned} n_2(\boldsymbol{r},\boldsymbol{r}') &= \langle\,\Psi_1\,|\,\psi^\dagger(\boldsymbol{r})\,\psi^\dagger(\boldsymbol{r}')\,\psi(\boldsymbol{r}')\,\psi(\boldsymbol{r})\,|\,\Psi_1\,\rangle \\ &= N(N-1)\int d^2r_3\cdots\int d^2r_N\,\big|\Psi_1(\boldsymbol{r},\boldsymbol{r}',\boldsymbol{r}_3,\ldots,\boldsymbol{r}_N)\big|^2 \\ &= \sum_{m_1}\sum_{m_2}\sum_{m_3}\sum_{m_4}\varphi^*_{m_1}(\boldsymbol{r})\,\varphi^*_{m_2}(\boldsymbol{r}')\,\varphi_{m_3}(\boldsymbol{r}')\,\varphi_{m_4}(\boldsymbol{r})\,\langle\,\Psi_1\,|\,c^\dagger_{m_1}\,c^\dagger_{m_2}\,c_{m_3}\,c_{m_4}\,|\,\Psi_1\,\rangle \end{aligned} \tag{3.27}$$

Now

$$\langle\,\Psi_1\,|\,c^\dagger_{m_1}\,c^\dagger_{m_2}\,c_{m_3}\,c_{m_4}\,|\,\Psi_2\,\rangle = \delta_{m_1,m_4}\,\delta_{m_2,m_3} - \delta_{m_1,m_3}\,\delta_{m_2,m_4} \tag{3.28}$$

and therefore

$$\begin{aligned} n_2(\boldsymbol{r},\boldsymbol{r}') &= \sum_m\sum_{m'}\Big(\big|\varphi_m(\boldsymbol{r})\big|^2\,\big|\varphi_{m'}(\boldsymbol{r}')\big|^2 - \varphi^*_m(\boldsymbol{r})\,\varphi_m(\boldsymbol{r}')\,\varphi^*_{m'}(\boldsymbol{r}')\,\varphi_{m'}(\boldsymbol{r})\Big) \\ &= \frac{1}{(2\pi\ell^2)^2}\Big(1 - e^{-(\boldsymbol{r}-\boldsymbol{r}')^2/2\ell^2}\Big) \equiv n_0^2\,g\big(|\boldsymbol{r}-\boldsymbol{r}'|\big) \equiv n_0^2\Big(1 + h\big(|\boldsymbol{r}-\boldsymbol{r}'|\big)\Big) \quad , \end{aligned} \tag{3.29}$$

where we have taken the $N_\phi \to \infty$ limit. Here $n_0 = \nu/2\pi\ell^2$ is the droplet density for the filled LL ($\nu = 1$), $g(r) = 1 - \exp(-r^2/2\ell^2)$ is the *pair distribution function* and

$$h(r) = g(r) - 1 = -\exp(-r^2/2\ell^2) \tag{3.30}$$

is the *pair correlation function.* The Coulomb energy per particle, once a neutralizing background is introduced, is given by

$$\frac{\langle V\rangle_{\rm corr}}{N} = \tfrac{1}{2}n\int d^2r\,v(r)\,h(r) = -\sqrt{\frac{\pi}{8}}\,\frac{e^2}{\epsilon\ell} \quad . \tag{3.31}$$

3.1.3 LLL projection

Consider the matrix element of a function $V(\boldsymbol{r})$ between two LLL states, $f(z)\ \exp(-z\bar{z}/4\ell^2)$ and $g(z)\ \exp(-z\bar{z}/4\ell^2)$. We define

$$\langle\, g \,|\, V \,|\, f \,\rangle = \int d^2r\ \overline{g(z)}\, V(\boldsymbol{r})\, f(z)\ \exp(-z\bar{z}/2\ell^2) \quad . \tag{3.32}$$

With $f(z) = \sum_{m=0}^{\infty} f_m\, z^m$ and $g(z) = \sum_{m=0}^{\infty} g_m\, z^m$, we have $\overline{g(z)} = \sum_{m=0}^{\infty} \bar{g}_m\, \bar{z}^m$, *i.e.* $\overline{g(z)} = \bar{g}(\bar{z})$. Now define the *normal ordered* operator $:V(\bar{z}, z):$ to be the function $V(\boldsymbol{r})$ expressed in terms of z and $\bar{z}$, but with all $\bar{z}$ factors to the left of all z factors. Thus, $:\boldsymbol{r}^2: \ = \bar{z}z$. When z and $\bar{z}$ commute, normal ordering accomplishes nothing. But note that

$$\langle\, g \,|\, V \,|\, f \,\rangle = \int d^2r\ \exp(-z\bar{z}/2\ell^2)\, \bar{g}(\bar{z})\, :V(2\ell^2\partial, z):\ f(z) \tag{3.33}$$

because we can integrate by parts, acting with $-2\ell^2\partial$ to the left, where it has no effect on $\bar{g}(\bar{z})$, which his holomorphic in $\bar{z}$, and which acts on the exponential factor as

$$-\,2\ell^2\partial\ \exp(-z\bar{z}/2\ell^2) = \bar{z}\ \exp(-z\bar{z}/2\ell^2) \quad , \tag{3.34}$$

thereby bringing down one factor of $\bar{z}$ for each application of $2\ell^2\partial$. Thus, the action of an operator $V(\boldsymbol{r})$ on the LLL wavefunction $\psi(\boldsymbol{r}) = f(z)\ \exp(-z\bar{z}/4\ell^2)$ is tantamount to acting only on the holomorphic part $f(z)$ with the operator[1] $:V(2\ell^2\partial, z):$. Thus, the Schrödinger equation in the LLL, dropping the constant $\frac{1}{2}\hbar\omega_{\rm c}$ zero point cyclotron energy term, is

$$:V(2\ell^2\partial, z):\ f(z) = Ef(z) \quad . \tag{3.35}$$

As an example, consider the harmonic potential $V(\boldsymbol{r}) = \frac{1}{2}K\boldsymbol{r}^2$. Projected to the LLL, the eigenstates in this potential have holomorphic parts $f(z)$ which satisfy

$$2\ell^2\, \frac{\partial}{\partial z}\left[z\, f(z)\right] = E\, f(z) \quad . \tag{3.36}$$

Clearly the solutions are the angular momentum states, with $f_m(z) = C_m z^m$, where C_m is a normalization constant. The energy eigenvalues are then $E_m = (m+1)K\ell^2$.

A particularly important application for us will be that of the plane wave, for which

$$:\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}):\ = \exp(-ik\ell^2\partial)\ \exp(-i\bar{k}z/2) \quad . \tag{3.37}$$

Note further that

$$:\exp(-i\boldsymbol{k}\cdot\boldsymbol{r}):\ f(z) = \exp(-\bar{k}k\ell^2/2)\ \exp(-i\bar{k}z/2)\, f(z - ik\ell^2) \quad . \tag{3.38}$$

Thus, the holomorphic coordinate within the function $f(z)$ is displaced by $-ik\ell^2$.

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

3.2 The Wigner Crystal

Let's first consider the interacting 2DEG in a field, but in the absence of disorder. The Hamiltonian is

$$H = \frac{1}{2m^*}\sum_{i=1}^{N}\left(\boldsymbol{p}_i + \tfrac{e}{c}\boldsymbol{A}_i\right)^2 + \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j) \quad , \tag{3.39}$$

with $v(r) = e^2/\epsilon r$ is the *three-dimensional* Coulomb interaction. We reiterate that while the electrons in are confined to a 2DEG, they interact via the three-dimensional $1/r$ potential and not the two-dimensional $\ln(1/r)$ form. This is because the field lines between charges in the 2DEG are themselves not confined to the 2DEG, but exist throughout the three-dimensional host heterostructure. The static dielectric constant in GaAs is $\epsilon = 13.13$.

In heterojunctions, the electron number N is fixed by the density of dopants. At fixed area A, the number of fluxoids in a finite area A under uniform $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$ is $N_\phi = BA/\phi_0$, where the Dirac flux quantum is $\phi_0 = hc/e = 4.137 \times 10^5\,\mathrm{T}\cdot\mathrm{\AA}^2$. The Landau level filling fraction, $\nu = N/N_\phi$, may then be adjusted by varying the field strength B. In Si MOSFETs, the electron density is set by the gate voltage V_g and can be varied during an experiment, as can B. Thus there are two ways to change ν in a MOSFET.

Recall that in the LLL, the kinetic energy is quenched, hence $H_{\mathsf{LLL}} = \Pi_0\, H\, \Pi_0 = \frac{1}{2}N\hbar\omega_\mathrm{c} + \widetilde{V}$, where

$$\widetilde{V} = \Pi_0 \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j)\,\Pi_0 = \sum_{i<j} v(\boldsymbol{\mathcal{R}}_i - \boldsymbol{\mathcal{R}}_j) \quad , \tag{3.40}$$

where $\boldsymbol{\mathcal{R}}_i = (\mathcal{X}_i, \mathcal{Y}_i)$ are the guiding center coordinates for the i^{th} particle[2]. Recall that

$$\big[\mathcal{X}_i, \mathcal{Y}_j\big] = -i\ell^2\,\delta_{ij} \quad , \tag{3.41}$$

with $\big[\mathcal{X}_i, \mathcal{X}_j\big] = \big[\mathcal{Y}_i, \mathcal{Y}_j\big] = 0$. We may drop the constant $\frac{1}{2}N\hbar\omega_\mathrm{c}$ piece in H_{LLL} and take the LLL-projected Hamiltonian to be $\widetilde{V}$. Projecting onto a single Landau level ignores Landau level mixing effects. We expect this approximation is justified provided the typical Coulomb energy scale $e^2\sqrt{\pi n}/\epsilon$ is sufficiently smaller than the cyclotron energy gap $\hbar\omega_\mathrm{c}$, or the Zeeman gap $\zeta\hbar\omega_\mathrm{c}$, where $\zeta = g^*m^*/2m_\mathrm{e}$ (see §2.2.5). With $n = \nu/2\pi\ell^2$ and $\omega_\mathrm{c} = \hbar/m^*\ell^2$, this criterion, up to dimensionless factors of order unity, is given by

$$\sqrt{\nu}\,\ell \ll \sqrt{2}\,a^*_{\mathsf{B}} \quad , \tag{3.42}$$

where $a^*_{\mathsf{B}} = \epsilon\hbar^2/m^*e^2$ is the effective Bohr radius, which is large in GaAs, with $m^* = 0.067\,m_\mathrm{e}$ and $\epsilon = 13$, we obtain $a^*_{\mathsf{B}} = 104\,\mathrm{\AA}$, and with $\ell = 257\,\mathrm{\AA}\,\sqrt{B[T]}$, our criterion then becomes $\nu \ll 0.33\,B[T]$, which is reasonably satisfied within the LLL ($\nu \le 1$) for $B \gtrsim 10\,\mathrm{T}$. In fact, there

[2]Recall that there are some subtleties associated with the LLL projection, having to do with normal-ordering, as discussed in §1.3.2.

will always be some degree of LL mixing, and the issue is really whether the actual ground state $|\,\Psi_0\,\rangle$ is adiabatically connected to some model or trial state that is conveniently expressed solely within the LLL, *i.e.* what *phase of matter* is present.

3.2.1 Classical Wigner crystal

At $\nu = 1$ there is a unique state corresponding to a filled LLL, but for $N < N_\phi$, the number of possible many-body states is given by the size of the Slater determinant basis, which is

$$\Omega(N, N_\phi) = \binom{N_\phi}{N} \simeq e^{c(\nu)N_\phi} \quad , \tag{3.43}$$

where $c(\nu) = -\nu \ln \nu - (1-\nu)\ln(1-\nu)$. There are thus exponentially many states to consider, so we need some intuition or physical principle to help us choose among them. One thing we can do is throw up our hands and *ignore quantum mechanics*[3] and pretend that the components of $\boldsymbol{\mathcal{R}}_i$ commute. This is equivalent to considering the classical potential energy function

$$V(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j) \quad , \tag{3.44}$$

with $v(r) = e^2/\epsilon r$. To be more precise, we could place our N particles in a circular disk of radius Λ, along with a uniform neutralizing background. Without the background, the energy will diverge as $N \to \infty$ with no thermodynamic limit (*i.e.* the energy will scale as N^2 rather than as N), but the neutralizing background, which is of course physical, fixes this problem[4]. What is the ground state? The simplest guess would be that it is a crystal. In some cases this can be proven mathematically, such as the case of 'sticky disks' where $v(r) = +\infty$ for $r < a$, $v(r) = -1$ for $r = a$, and $v(r) = 0$ for $r > 0$. For this case, the ground state is a triangular Bravais lattice[5]. Recall that the Abrikosov vortex lattice in a type-II s-wave superconductor, where the vortices interact by a screened repulsive logarithmic potential, is also triangular. That *weak crystallization*, meaning crystallization in a weakly first-order transition, should result in a triangular lattice in $d = 2$ was argued by Alexander and McTague[6] based on a Landau theory of the transition. The argument is as follows. Let $\varrho_{\boldsymbol{G}}$ be the amplitude of the Fourier component of the density $\varrho(\boldsymbol{r})$ with wavevector $\boldsymbol{G}$, which is a reciprocal lattice vector of the

[3] I am a trained professional. Students should not try this at home.

[4] Recall that in heterostructures, the neutralizing background is due to the dopants, which are typically recessed by several hundred Ångstroms from the 2DEG.

[5] R. C. Heitmann and C. Radin, *J. Stat. Phys.* **22**, 281 (1980) and also C. Radin, *J. Stat. Phys.* **26**, 365 (1981), where the original proof was extended to 'soft disks' where there is an annulus over which the interaction potential varies linearly with distance between $v(a) = -1$ and $v(b) = 0$.

[6] S. Alexander and J. McTague, *Phys. Rev. Lett.* **41**, 702 (1978). See also E. I. Kats, V. V. Lebedev, and A. R. Muranov, *Phys. Rep.* **228**, 1 (1993).

incipient crystalline phase. Then construct the free energy

$$F\big[\{\varrho_{\boldsymbol{G}}\}\big] = \frac{1}{2}\sum_{\boldsymbol{G}} \chi^{-1}(\boldsymbol{G})\,|\varrho_{\boldsymbol{G}}|^2 - \frac{1}{3}\,B\sum_{\boldsymbol{G}_1}\sum_{\boldsymbol{G}_2}\sum_{\boldsymbol{G}_3} \varrho_{\boldsymbol{G}_1}\,\varrho_{\boldsymbol{G}_2}\,\varrho_{\boldsymbol{G}_3}\,\delta_{\boldsymbol{G}_1+\boldsymbol{G}_2+\boldsymbol{G}_3,\boldsymbol{0}} + \frac{1}{4}\,C\sum_{\boldsymbol{G}_1}\sum_{\boldsymbol{G}_2}\sum_{\boldsymbol{G}_3}\sum_{\boldsymbol{G}_4} \varrho_{\boldsymbol{G}_1}\,\varrho_{\boldsymbol{G}_2}\,\varrho_{\boldsymbol{G}_3}\,\varrho_{\boldsymbol{G}_4}\,\delta_{\boldsymbol{G}_1+\boldsymbol{G}_2+\boldsymbol{G}_3+\boldsymbol{G}_4,\boldsymbol{0}} + \cdots \quad , \tag{3.45}$$

where

$$\chi^{-1}(\boldsymbol{k}) = r + b\,(\boldsymbol{k}^2 - G^2)^2 \tag{3.46}$$

is the inverse static susceptibility at wavevector $\boldsymbol{k}$, which for fixed r is minimized for $|\boldsymbol{k}| = G$. The quadratic term determines the magnitude of the preferred wavevectors at which condensation takes place at $r = r_{\rm c} = 0$, but this energy is degenerate over the circle (or sphere in $d = 3$) of radius G. For weak crystallization, then, the cubic term determines the crystal structure, and evidently prefers structures whose reciprocal lattices contain the maximum number of triangles, in order to satisfy the $\boldsymbol{G}_1 + \boldsymbol{G}_2 + \boldsymbol{G}_3 = 0$ condition. In $d = 2$ this prefers a reciprocal lattice which is triangular, hence the underlying direct Bravais lattice is also triangular (or honeycomb). In $d = 3$, this condition prefers the fcc structure among all regular lattices, and the corresponding direct lattice is thus bcc. It should be emphasized that the Alexander-McTague theory applies to the weak crystallization of a fluid, and really describes the formation of a charge density wave structure, rather than a Wigner crystal of point particles.

The energy per particle of $d = 2$ crystalline lattices of charges interacting by the potential $v(r) = e^2/\epsilon r$, in the presence of a uniform neutralizing background, was computed by Bonsall and Maradudin (BM) using the Ewald summation method[7]. They obtained the general result,

$$u_{\rm WC} = \frac{U_{\rm WC}}{N} = -\frac{e^2}{\epsilon\sqrt{\Omega}}\left\{2 - {\sum_{\boldsymbol{R}}}' \phi_{-1/2}(\pi\boldsymbol{R}^2/\Omega)\right\} \tag{3.47}$$

where the sum is over all nonzero direct lattice vectors $\boldsymbol{R}$, Ω is the unit cell area, and

$$\phi_n(z) = \int_1^\infty dt\; t^n\, e^{-zt} \tag{3.48}$$

is known as the Misra function. BM obtained the following results:

$$u_{\rm WC} = -\frac{e^2}{\epsilon}\left(\frac{2\pi}{\Omega}\right)^{1/2} \times \begin{cases} 0.777990 & \text{(square)} \\ 0.782133 & \text{(triangular)} \end{cases} \quad . \tag{3.49}$$

Thus, the triangular lattice configuration has lower energy per particle. Note that $n\Omega = 1$ where $n = \nu/2\pi\ell^2$ is the density, hence $(2\pi/\Omega)^{1/2} = \sqrt{\nu}/\ell$.

[7] L. Bonsall and A. A. Maradudin, *Phys. Rev. B* **15**, 1959 (1977).

3.2.2 Quantum Wigner crystal

How can we restore quantum mechanics, *i.e.* the noncommutativity of $[\mathcal{X}_i, \mathcal{Y}_j] = -i\ell^2\,\delta_{ij}$? Maki and Zotos[8] constructed a trial LLL Wigner crystal wavefunction, using the coherent states $\varphi_{\boldsymbol{R}}(\boldsymbol{r})$ as a basis, where the set $\{\boldsymbol{R}\}$ corresponds to a triangular lattice, *i.e.* $\boldsymbol{R}_{mn} = m\boldsymbol{a}_1 + n\boldsymbol{a}_2$ with $\boldsymbol{a}_{1,2} = \frac{1}{2}a\big(\hat{\boldsymbol{x}} \mp \sqrt{3}\,\hat{\boldsymbol{y}}\big)$ and $\Omega = n^{-1} = \frac{1}{2}\sqrt{3}\,a^2$. Recall the form of the LLL coherent state wavefunction from §1.3.7,

$$\varphi_{\boldsymbol{R}}(\boldsymbol{r}) = \langle\, \boldsymbol{r} \,|\, \boldsymbol{R} \,\rangle = (2\pi\ell^2)^{-1/2} \exp\big(i\boldsymbol{R}\times\boldsymbol{r}\cdot\hat{\boldsymbol{z}}/2\ell^2\big) \exp\big(-|\boldsymbol{r}-\boldsymbol{R}|^2/4\ell^2\big) \quad . \tag{3.50}$$

The Maki-Zotos wavefunction is then

$$|\,\Psi_{\rm MZ}\,\rangle = N^{-1/2} \sum_{\sigma\in\mathcal{S}_N} \mathsf{sgn}(\sigma)\,|\,\boldsymbol{R}_{\sigma(1)},\ldots,\boldsymbol{R}_{\sigma(N)}\,\rangle \quad . \tag{3.51}$$

Note that the MZ wavefunction is not normalized, due to the fact that the von Neumann lattice of coherent states is not an orthonormal basis:

$$\langle\,\Psi_{\rm MZ}\,|\,\Psi_{\rm MZ}\,\rangle = \sum_{\sigma\in\mathcal{S}_N}\prod_{j=1}^{N}\langle\,\boldsymbol{R}_j\,|\,\boldsymbol{R}_{\sigma(j)}\,\rangle \quad , \tag{3.52}$$

where

$$\langle\,\boldsymbol{R}\,|\,\boldsymbol{R}'\,\rangle = \exp\big(-i\boldsymbol{R}\times\boldsymbol{R}'\cdot\hat{\boldsymbol{z}}/2\ell^2\big)\exp\big(-|\boldsymbol{R}-\boldsymbol{R}'|^2/4\ell^2\big) \quad . \tag{3.53}$$

Therefore, in computing the expectation value of any operator $\mathcal{O}$ in the MZ state, one must compute

$$\langle\mathcal{O}\rangle_{\rm MZ} = \frac{\langle\,\Psi_{\rm MZ}\,|\,\mathcal{O}\,|\,\Psi_{\rm MZ}\,\rangle}{\langle\,\Psi_{\rm MZ}\,|\,\Psi_{\rm MZ}\,\rangle} \quad . \tag{3.54}$$

For nearest neighbors, the overlap magnitude is $\big|\langle\,\boldsymbol{R}\,|\,\boldsymbol{R}'\,\rangle\big| = \exp(-a^2/4\ell^2) = \exp(2\pi/\sqrt{3}\,\nu)$, which says that exchange effects are negligible in the low density limit $\nu \to 0$.

Taking into account only electron-electron interactions, *i.e.* with no neutralizing background as of yet, the energy of the MZ Wigner crystal atate is

$$E = \frac{e^2}{\epsilon}\,\frac{\langle\,\Psi_{\rm MZ}\,|\,\sum_{i<j}\frac{1}{r_{ij}}\,|\,\Psi_{\rm MZ}\,\rangle}{\langle\,\Psi_{\rm MZ}\,|\,\Psi_{\rm MZ}\,\rangle} = \sum_{i<j}\widetilde{V}_2(\boldsymbol{R}_i-\boldsymbol{R}_j) + \sum_{i<j<k}\widetilde{V}_3(\boldsymbol{R}_i,\boldsymbol{R}_j,\boldsymbol{R}_k) + \ldots \quad , \tag{3.55}$$

where

$$\begin{aligned}\widetilde{V}_2(\boldsymbol{R}) &= \frac{e^2}{\epsilon}\,\frac{\langle\,\boldsymbol{0},\boldsymbol{R}\,|\,\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\,|\,\boldsymbol{0},\boldsymbol{R}\,\rangle - \langle\,\boldsymbol{0},\boldsymbol{R}\,|\,\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\,|\,\boldsymbol{R},\boldsymbol{0}\,\rangle}{\langle\,\boldsymbol{0},\boldsymbol{R}\,|\,\boldsymbol{0},\boldsymbol{R}\,\rangle - \langle\,\boldsymbol{0},\boldsymbol{R}\,|\,\boldsymbol{R},\boldsymbol{0}\,\rangle} \\ &= \frac{\sqrt{\pi}\,e^2}{4\epsilon\ell}\,\mathrm{sech}(R^2/8\ell^2)\,I_0(R^2/8\ell^2) = \begin{cases}\sqrt{\pi}\,e^2/4\epsilon\ell & R\to 0\\ e^2/\epsilon R & R\to\infty\end{cases} \quad . \\ &= \frac{e^2}{\epsilon R}\left\{1+\frac{\ell^2}{R^2}+\frac{9\ell^4}{2R^4}+\frac{75\ell^6}{2R^6}+\ldots\right\} \quad .\end{aligned} \tag{3.56}$$

[8] K. Maki and X. Zotos, *Phys. Rev. B* **28**, 4849 (1983).

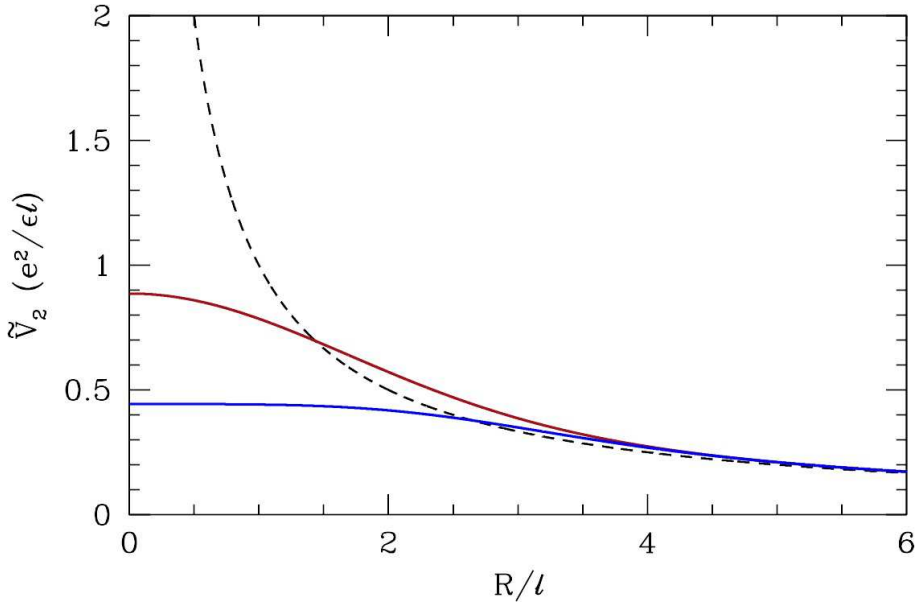


Figure 3.1: Bare (dashed, black), projected (red) and exchange-corrected (blue) Coulomb interaction in the LLL.

For details on the three-body term, see Maki and Zotos. Note that without the exchange correction, we would have

$$\begin{aligned}\widetilde{V}_2'(\boldsymbol{R}) &= \frac{e^2}{\epsilon}\,\langle\,\mathbf{0},\boldsymbol{R}\,|\,\frac{1}{|\boldsymbol{r}-\boldsymbol{r}'|}\,|\,\mathbf{0},\boldsymbol{R}\,\rangle \\ &= \frac{\sqrt{\pi}\,e^2}{2\epsilon\ell}\,\exp(-R^2/8\ell^2)\,I_0(R^2/8\ell^2) = \begin{cases}\sqrt{\pi}\,e^2/2\epsilon\ell & R\to 0\\ e^2/\epsilon R & R\to\infty\end{cases}\quad.\end{aligned} \tag{3.57}$$

The consequences of projection and exchange correction are shown in Fig. 3.1.

The first term in the expansion of Eqn. 3.56 in powers of R^{-1} gives the classical energy, which diverges as N^2 in the absence of a neutralizing background. With the background, this term gives us the Bonsall-Maradudin result $u_{\rm WC} = -0.782133\,\sqrt{\nu}\,e^2/\epsilon\ell$ per particle. The remaining contributions, which are positive, are the quantum contributions to the correlation energy per particle. Note that the correlation energy of the filled Landau level, which we computed in Eqn. 3.31, is $u_{\rm corr}(\nu=1) = -0.626657\,e^2/\epsilon\ell$, which is clearly greater than the classical WC energy at this density. This is due to zero point quantum fluctuations of the electron coordinates relative to their classical energy-minimizing locations. Maki and Zotos found that their correlation energy compared well with Hartree-Fock CDW calculations by Yoshioka and Lee[9], with agreement to 1% throughout the regime $\nu < \frac{1}{2}$. Exchange effects were important to consider for $\nu \in [0.4, 0.5]$ but accounted for less than a percent of the correlation energy at lower fillings. The salient feature here is that the correlation energy $u_{\rm WC}(\nu)$ is a *smooth* function of the filling ν.

[9]D. Yoshioka and P. A. Lee, *Phys. Rev. B* **27**, 4986 (1983).

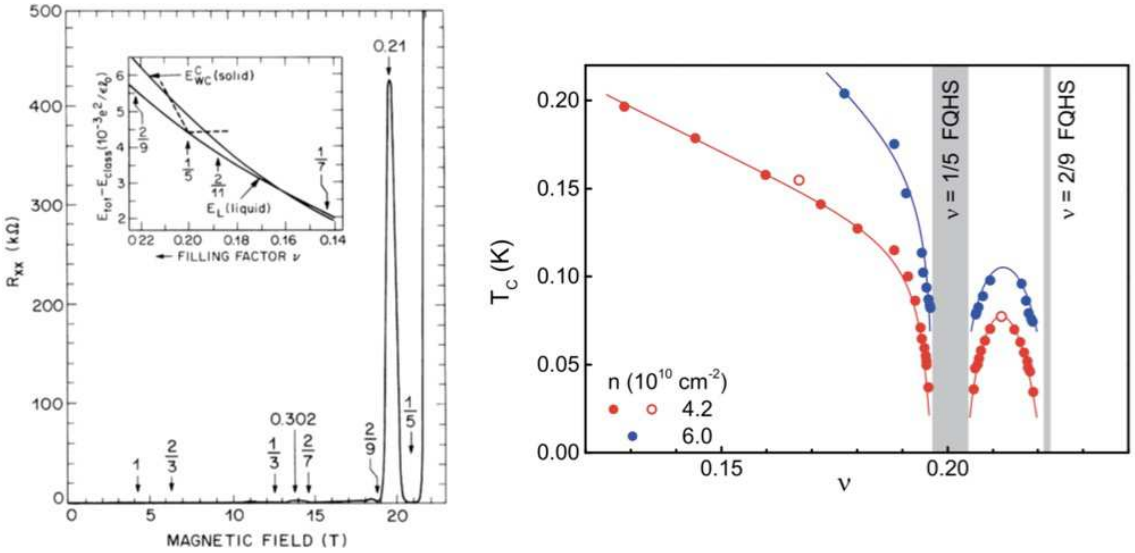


Figure 3.2: Evidence of Wigner crystal behavior for $\nu \lesssim 0.2$. Left: Data from H. W. Jiang *et al.*, *Phys. Rev. Lett.* **65**, 633 (1990). The longitudinal resistance R_{xx} exhibits pronounced peaks, dwarfing those in all FQH states, for $\nu \approx 0.21$ and for $\nu < 0.19$, suggesting reentrant solid behavior. Right: Wigner crystal phase diagram inferred from capacitive measurements of effective screening by H. Deng *et al.*, *Phys. Rev. Lett.* **122**, 116601 (2019), showing reentrant solid behavior between $\nu = \frac{1}{5}$ and $\nu = \frac{2}{9}$. The WC phase screens very poorly, and screening efficiency improves once the WC melts.

As we shall see, this is inconsistent with the phenomenology of the FQHE, which requires that the free energy $F(\nu)$ have *cusps* when ν corresponds to the filling at a FQH plateau. In addition, a Wigner crystal or charge density wave state breaks translational invariance, and is subject to pinning and the formation of Imry-Ma domains[10]. Observations of Wigner crystal phases of GaAs/AlGaAs heterojunction 2DEGs were first reported by E. Andrei *et al.*, *Phys. Rev. Lett.* **60**, 2765 (1988).

3.2.3 Magnetophonons in the Wigner crystal and in charged elastic media

Within the harmonic approximation, the energy of the deformed crystal, with $\boldsymbol{R} \to \boldsymbol{R} + \boldsymbol{u}_{\boldsymbol{R}}$, is

$$\begin{aligned} U &= \frac{1}{2} \sum_{\boldsymbol{R} \neq \boldsymbol{R}'} \widetilde{V}\big(\boldsymbol{R} + \boldsymbol{u}_{\boldsymbol{R}} - \boldsymbol{R}' - \boldsymbol{u}_{\boldsymbol{R}'}\big) \\ &= U_0 + \frac{1}{2} \sum_{\boldsymbol{R} \neq \boldsymbol{R}'} \big(u^\alpha_{\boldsymbol{R}} - u^\alpha_{\boldsymbol{R}'}\big)\big(u^\beta_{\boldsymbol{R}} - u^\beta_{\boldsymbol{R}'}\big) \, \frac{\partial^2 \widetilde{V}(\boldsymbol{R} - \boldsymbol{R}')}{\partial R^\alpha \, \partial R^\beta} + \dots \quad , \end{aligned} \tag{3.58}$$

[10] H. Fukuyama and P. A. Lee, *Phys. Rev. B* **18**, 6245 (1978).

where each $\boldsymbol{R}$ is a triangular lattice site. Now write

$$u^\alpha_{\boldsymbol{R}} = \frac{1}{\sqrt{N_{\rm c}}} \sum_{\boldsymbol{k}} \hat{u}^\alpha(\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot\boldsymbol{R}} \tag{3.59}$$

where $N_{\rm c} = N$ is the number of unit cells[11], in which case

$$U = U_0 + \frac{1}{2} \sum_{\boldsymbol{k}} \sum_{\alpha,\beta} \hat{\Phi}_{\alpha\beta}(\boldsymbol{k})\, \hat{u}^\alpha(\boldsymbol{k})\, \hat{u}^\beta(-\boldsymbol{k}) + \mathcal{O}(u^3) \quad , \tag{3.60}$$

with the dynamical matrix

$$\hat{\Phi}_{\alpha\beta}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} (1 - \cos \boldsymbol{k}\cdot\boldsymbol{R})\, \frac{\partial^2 \widetilde{V}(\boldsymbol{R})}{\partial R^\alpha\, \partial R^\beta} \quad . \tag{3.61}$$

We now write $\widetilde{V}(\boldsymbol{R})$ as a Fourier integral, *viz.*

$$\widetilde{V}(\boldsymbol{R}) = \int \frac{d^2q}{(2\pi)^2}\, \hat{\widetilde{V}}(\boldsymbol{q})\, e^{i\boldsymbol{q}\cdot\boldsymbol{R}} \quad , \tag{3.62}$$

and make use of the Poisson summation formula,

$$\sum_{\boldsymbol{R}} e^{i\boldsymbol{k}\cdot\boldsymbol{R}} = \frac{(2\pi)^2}{\Omega} \sum_{\boldsymbol{G}} \delta(\boldsymbol{k} - \boldsymbol{G}) \quad , \tag{3.63}$$

where $\Omega = 1/n$ is the area per unit cell. This gives

$$\hat{\Phi}_{\alpha\beta}(\boldsymbol{k}) = \frac{\nu}{2\pi\ell^2} \sum_{\boldsymbol{G}} \Big[(G_\alpha + k_\alpha)(G_\beta + k_\beta)\, \hat{\widetilde{V}}(\boldsymbol{G} + \boldsymbol{k}) - G_\alpha G_\beta\, \hat{\widetilde{V}}(\boldsymbol{G}) \Big] \tag{3.64}$$

Now we may quantize, writing

$$\hat{z}(\boldsymbol{k}) = \hat{u}_x(\boldsymbol{k}) + i\hat{u}_y(\boldsymbol{k}) = \sqrt{2}\, \ell\, b^\dagger_{-\boldsymbol{k}} \quad , \tag{3.65}$$

in the LLL (*cf.* Eqn. 1.48), to obtain the magnetophonon Hamiltonian,

$$H^0_{\rm MP} = \sum_{\boldsymbol{k}} \Big[\Omega_{\boldsymbol{k}} \big(b^\dagger_{\boldsymbol{k}}\, b_{\boldsymbol{k}} + b^\dagger_{-\boldsymbol{k}}\, b_{-\boldsymbol{k}} \big) + \Delta_{\boldsymbol{k}}\, b_{\boldsymbol{k}}\, b_{-\boldsymbol{k}} + \Delta^*_{\boldsymbol{k}}\, b^\dagger_{\boldsymbol{k}}\, b^\dagger_{-\boldsymbol{k}} \Big] \quad , \tag{3.66}$$

where

$$\begin{aligned} \Omega_{\boldsymbol{k}} &= \tfrac{1}{2}\ell^2 \left[\hat{\Phi}_{xx}(\boldsymbol{k}) + \hat{\Phi}_{yy}(\boldsymbol{k}) \right] \\ \Delta_{\boldsymbol{k}} &= \tfrac{1}{2}\ell^2 \left[\hat{\Phi}_{xx}(\boldsymbol{k}) - \hat{\Phi}_{yy}(\boldsymbol{k}) + 2i\hat{\Phi}_{xy}(\boldsymbol{k}) \right] \quad . \end{aligned} \tag{3.67}$$

[11] There is one electron per unit cell, because the triangular lattice is a Bravais lattice.

Diagonalizing $H^0_{\rm MP}$ via a Bogoliubov transformation, we obtain the dispersion

$$\omega_{\boldsymbol{k}} = \sqrt{\Omega^2_{\boldsymbol{k}} - |\Delta_{\boldsymbol{k}}|^2} = \ell \sqrt{\hat\Phi_{xx}(\boldsymbol{k})\,\hat\Phi_{yy}(\boldsymbol{k}) - \hat\Phi^2_{xy}(\boldsymbol{k})} \quad . \tag{3.68}$$

Let $\hat{\widetilde{V}}(\boldsymbol{k}) = (2\pi e^2\ell/\epsilon)\hat F(k\ell)$. Then

$$\hat\Phi_{\alpha\beta}(\boldsymbol{k}) = \frac{\nu e^2}{\epsilon\ell}\Big(\big[\hat F(k\ell) + C_0\big]k_\alpha\,k_\beta + C_1\boldsymbol{k}^2\,\delta_{\alpha\beta} + \mathcal{O}(k^3)\Big) \tag{3.69}$$

with

$$\begin{aligned} C_0 &= {\sum_{\boldsymbol{G}}}' \big[\hat F(G\ell) + \tfrac{7}{8}\,(G\ell)\hat F'(G\ell) + \tfrac{1}{8}\,(G\ell)^2\hat F''(G\ell)\big] \\ C_1 &= {\sum_{\boldsymbol{G}}}' \big[\tfrac{3}{16}\,(G\ell)\hat F'(G\ell) + \tfrac{1}{16}\,(G\ell)^2\hat F''(G\ell)\big] \quad . \end{aligned} \tag{3.70}$$

where the primes on the sums indicate that $\boldsymbol{G} = 0$ is excluded. For the unprojected Coulomb potential $v(r) = e^2/\epsilon r$, the Fourier transform yields $\hat F(k\ell) = 1/k\ell$ and the sums fail to converge. One must then reformulate the problem using the Ewald summation method. However in our case, $\widetilde V(\boldsymbol{r})$ is the LLL-projected and exchange-corrected potential of Eqn. 3.56. In this case $\hat F(k\ell)$ behaves as $1/k\ell$ in the infrared (*i.e.* as $k \to 0$), but in the ultraviolet the short distance blunting of the $1/r$ divergence from the LLL projection results in an exponential decay in $k\ell$, as in the case of the Yukawa potential. In this case, the sums for $C_{0,1}$ converge nicely[12]. It is left as an exercise to the reader to verify the long wavelength dispersion,

$$\omega_{\boldsymbol{k}} = \frac{\nu e^2}{\epsilon}\,k^2\sqrt{C_0\,\hat F(k\ell) + C_1(C_0 + C_1)} \quad . \tag{3.71}$$

As $k \to 0$, then, we have $F(k\ell) = 1/k\ell$ dominates inside the radical, and $\omega_{\boldsymbol{k}} \propto k^{3/2}$. Note that if $F(k\ell)$ were to approach a constant as $k \to 0$, corresponding to $v(\boldsymbol{r}) \sim \delta(\boldsymbol{r})$, we'd have $\omega_{\boldsymbol{k}} \propto k^2$. Conversely, if the potential were logarithmic, we'd obtain $\omega_{\boldsymbol{k}} \propto k$.

Classical derivation

Consider an elastic medium with potential energy density

$$\mathcal{U}(\boldsymbol{x}) = \mu\,\mathsf{Tr}(\varepsilon^2) + \tfrac{1}{2}\,\lambda\,(\mathsf{Tr}\;\varepsilon)^2 \tag{3.72}$$

where $\varepsilon(\boldsymbol{x})$ is the symmetric strain tensor, with components

$$\varepsilon_{\alpha\beta} = \frac{1}{2}\left(\frac{\partial u^\alpha}{\partial x^\beta} + \frac{\partial u^\beta}{\partial x^\alpha}\right) \quad , \tag{3.73}$$

[12] We assume both C_0 and C_1 are positive.

where $\boldsymbol{u}(\boldsymbol{x})$ is the local displacement field.. The Lagrangian density is

$$\mathcal{L} = \tfrac{1}{2}\, n_0\, m\big(\dot{u}_x^2 + \dot{u}_y^2\big) + \frac{neB}{2c}\,\big(u_x\,\dot{u}_y - u_y\,\dot{u}_x\big) - \mathcal{U}(\boldsymbol{x}) \quad , \tag{3.74}$$

where n_0 is the number density, $n_0\, m$ is the mass density and $(-n_0\, e)$ is the charge density. Writing the action in terms of the Fourier modes $u^\alpha_{\boldsymbol{k}}$, we have

$$S = \int dt \sum_{\boldsymbol{k}} \left\{ \tfrac{1}{2} n_0\, m\, \dot{u}^\alpha_{\boldsymbol{k}}\, \dot{u}^\alpha_{-\boldsymbol{k}} + \frac{n_0\, eB}{2c}\, \epsilon_{\alpha\beta}\, u^\alpha_{\boldsymbol{k}}\, \dot{u}^\beta_{-\boldsymbol{k}} - \tfrac{1}{2}\Big[\mu(\delta^{\alpha\beta} - \hat{k}^\alpha \hat{k}^\beta) + (\lambda + 2\mu)\, \hat{k}^\alpha \hat{k}^\beta\Big] \boldsymbol{k}^2 \right\} \quad . \tag{3.75}$$

We now express $u^\alpha_{\boldsymbol{k}}$ in terms of longitudinal and transverse modes:

$$\boldsymbol{u}_{\boldsymbol{k}} = i\hat{\boldsymbol{k}}\, u^{\parallel}_{\boldsymbol{k}} + i\hat{\boldsymbol{z}} \times \hat{\boldsymbol{k}}\, u^{\perp}_{\boldsymbol{k}} \quad . \tag{3.76}$$

The factors of i ensure that $\boldsymbol{u}^*_{\boldsymbol{k}} = \boldsymbol{u}_{-\boldsymbol{k}}$ if $(u^{\parallel/\perp}_{\boldsymbol{k}})^* = u^{\parallel/\perp}_{-\boldsymbol{k}}$, *i.e.* they are all Fourier components of *real* fields. Now we have $L = T - U$ with

$$\begin{aligned} T &= \sum_{\boldsymbol{k}} \left\{ \tfrac{1}{2} n_0\, m\big(\dot{u}^{\parallel}_{\boldsymbol{k}}\, \dot{u}^{\parallel}_{-\boldsymbol{k}} + \dot{u}^{\perp}_{\boldsymbol{k}}\, \dot{u}^{\perp}_{-\boldsymbol{k}}\big) + \frac{n_0\, eB}{2c}\, \big(u^{\parallel}_{\boldsymbol{k}}\, \dot{u}^{\perp}_{-\boldsymbol{k}} - u^{\perp}_{\boldsymbol{k}}\, \dot{u}^{\parallel}_{-\boldsymbol{k}}\big) \right\} \\ U &= \sum_{\boldsymbol{k}} \left\{ \tfrac{1}{2}(\lambda + 2\mu)\, \boldsymbol{k}^2 \left|u^{\parallel}_{\boldsymbol{k}}\right|^2 + \tfrac{1}{2}\mu\, \boldsymbol{k}^2 \left|u^{\perp}_{\boldsymbol{k}}\right|^2 \right\} \quad . \end{aligned} \tag{3.77}$$

Now write the Lagrangian $L = T - U$ and take the functional variation of the action $S = \int dt\, L$ with respect to $u^{\parallel}_{-\boldsymbol{k}}$ and with respect to $u^{\perp}_{-\boldsymbol{k}}$ to get

$$\begin{aligned} \frac{\delta S}{\delta u^{\parallel}_{-\boldsymbol{k}}} = 0 \quad &\Rightarrow \quad n_0\, m\, \ddot{u}^{\parallel}_{\boldsymbol{k}} - \frac{n_0\, eB}{c}\, \dot{u}^{\perp}_{\boldsymbol{k}} = (\lambda + 2\mu)\, \boldsymbol{k}^2\, u^{\parallel}_{\boldsymbol{k}} \\ \frac{\delta S}{\delta u^{\perp}_{-\boldsymbol{k}}} = 0 \quad &\Rightarrow \quad n_0\, m\, \ddot{u}^{\perp}_{\boldsymbol{k}} + \frac{n_0\, eB}{c}\, \dot{u}^{\parallel}_{\boldsymbol{k}} = \mu\, \boldsymbol{k}^2\, u^{\perp}_{\boldsymbol{k}} \quad . \end{aligned} \tag{3.78}$$

In frequency space, this is equivalent to the system

$$\begin{pmatrix} \omega^2 - \omega^2_{\mathsf{L}}(\boldsymbol{k}) & i\,\omega\,\omega_{\mathrm{c}} \\ -i\,\omega\,\omega_{\mathrm{c}} & \omega^2 - \omega^2_{\mathsf{T}}(\boldsymbol{k}) \end{pmatrix} \begin{pmatrix} u^{\parallel}_{\boldsymbol{k}} \\ u^{\perp}_{\boldsymbol{k}} \end{pmatrix} = 0 \quad , \tag{3.79}$$

where $\omega_{\mathrm{c}} = eB/mc$, and where

$$\omega_{\mathsf{L}}(\boldsymbol{k}) = \left(\frac{\lambda + 2\mu}{n_0\, m}\right)^{1/2} |\boldsymbol{k}| \qquad , \qquad \omega_{\mathsf{T}}(\boldsymbol{k}) = \left(\frac{\mu}{n_0\, m}\right)^{1/2} |\boldsymbol{k}| \tag{3.80}$$

are the long wavelength longitudinal and transverse phonon dispersions when $B = 0$. Setting the determinant to zero, we obtain the two normal modes,

$$\omega_\pm(\boldsymbol{k}) = \left[\tfrac{1}{2}\big[\omega^2_{\mathrm{c}} + \omega^2_{\mathsf{L}}(\boldsymbol{k}) + \omega^2_{\mathsf{T}}(\boldsymbol{k})\big]^2 \pm \tfrac{1}{2}\sqrt{\big(\omega^2_{\mathrm{c}} + \omega^2_{\mathsf{L}}(\boldsymbol{k}) + \omega^2_{\mathsf{T}}(\boldsymbol{k})\big)^2 - 4\,\omega^2_{\mathsf{L}}(\boldsymbol{k})\,\omega^2_{\mathsf{T}}(\boldsymbol{k})} \right]^{1/2} \quad . \tag{3.81}$$

In the long wavelength ($k \to 0$) limit, then,

$$\begin{aligned}\omega_+(\boldsymbol{k}) &= \omega_{\rm c} + \frac{\omega_{\sf L}^2(\boldsymbol{k}) + \omega_{\sf T}^2(\boldsymbol{k})}{2\omega_{\rm c}} + \dots \\ \omega_-(\boldsymbol{k}) &= \frac{\omega_{\sf L}(\boldsymbol{k})\,\omega_{\sf T}(\boldsymbol{k})}{\omega_{\rm c}} + \dots \quad .\end{aligned} \tag{3.82}$$

Since both $\omega_{\sf L}(\boldsymbol{k})$ and $\omega_{\sf T}(\boldsymbol{k})$ vanish linearly with k in this limit, we find that the lower mode disperses as k^2 and the upper mode is gapped with $\omega_+(0) = \omega_{\rm c}$.

How do we add Coulomb interactions to this model? Note that $\mathsf{Tr}\,\varepsilon = \boldsymbol{\nabla}\cdot\boldsymbol{u}$, which is related to the local variation of the number density according to

$$n(\boldsymbol{x}) = n_0\left(1 + \boldsymbol{\nabla}\cdot\boldsymbol{u}\right) \quad , \tag{3.83}$$

i.e. $\delta n(\boldsymbol{x}) = n(\boldsymbol{x}) - n_0(\boldsymbol{x}) = n_0\,\boldsymbol{\nabla}\cdot\boldsymbol{u}$. Thus

$$\Delta U = \frac{1}{2}\int d^2x \int d^2x'\;\delta n(\boldsymbol{x})\,v(\boldsymbol{x}-\boldsymbol{x}')\,\delta n(\boldsymbol{x}') = \frac{n_0^2}{2}\sum_{\boldsymbol{k}}\hat{v}(\boldsymbol{k})\,\boldsymbol{k}^2\left|u_{\boldsymbol{k}}^{\parallel}\right|^2 \quad , \tag{3.84}$$

and the Coulomb interaction $\hat{v}(\boldsymbol{k}) = 2\pi e^2/\epsilon|\boldsymbol{k}|$ is accommodated by the replacement of the Lamé parameter λ with an effective Lamé parameter $\lambda(\boldsymbol{k})$, *viz.*

$$\lambda \to \lambda(\boldsymbol{k}) = \lambda + n_0^2\,\hat{v}(\boldsymbol{k}) \quad . \tag{3.85}$$

We then have $\omega_{\sf L}(\boldsymbol{k}) = (2\pi n_0\,e^2/\epsilon m)^{1/2}\,k^{1/2}$ at long wavelengths. This is the dispersion of the two-dimensional plasmon. Note that for $\hat{v}(\boldsymbol{k}) \propto k^{-2}$, as would be the case for a logarithmic potential, the 2D plasmon would be gapped, as is the plasmon in $d = 3$ with $1/r$ interactions, corresponding to $\hat{v}_{\rm 3D}(\boldsymbol{k}) = 4\pi e^2/\epsilon\boldsymbol{k}^2$. The $k^{1/2}$ longitudinal mode for $B = 0$ in $d = 2$ then entails $\omega_-(\boldsymbol{k}) \propto k^{3/2}$ for a charged elastic medium in a uniform magnetic field. This is the famous $k^{3/2}$ magnetophonon!

3.2.4 Imry-Ma argument: pinning by quenched disorder

In a crystalline phase there is long-ranged positional order and a breaking of the continuous symmetry of translation. It then behooves us to ask how such phases fare in the presence of quenched disorder, which in the case of QH systems is due to the random positions of dopant ions, each of which becomes a Coulomb impurity scatterer. The issue of how quenched randomness affects a system's attempt to order was taken up in a beautiful paper by Imry and Ma in 1975[13]. Quenched disorder in these systems is typically modeled as a local field. In systems with discrete symmetries, such as the Ising model, one would take $V_{\sf dis} = -\sum_{\boldsymbol{r}} H_{\boldsymbol{r}}\,\sigma_{\boldsymbol{r}}$. In

[13] J. Imry and S.-K. Ma, *Phys. Rev. Lett.* **35**, 1399 (1975).

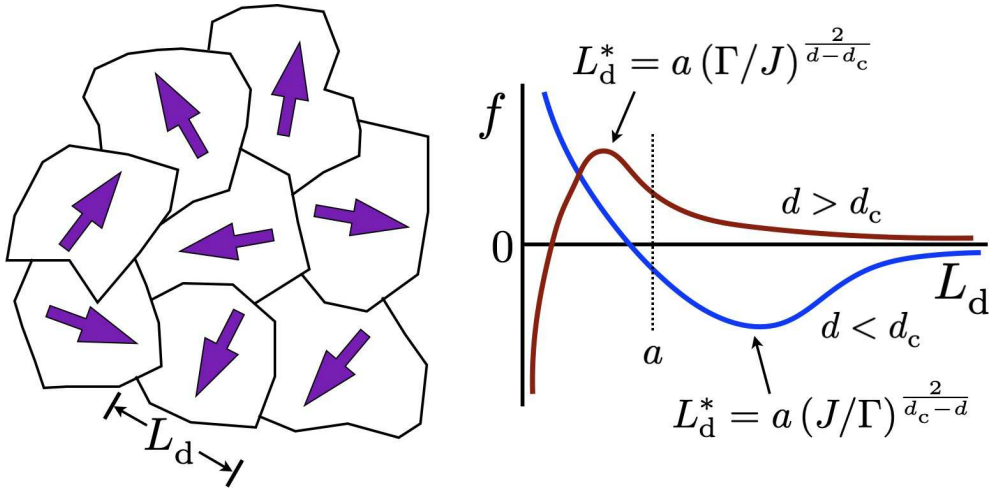


Figure 3.3: Adding disorder to systems which spontaneously break a discrete or continuous symmetry results in the formation of Imry-Ma domains of finite size $L_{\rm d}$ if d is less than the critical dimension $d_{\rm c}$, where $d_{\rm c} = 2$ for discrete symmetry and $d_{\rm c} = 4$ for continuous symmetry. Left: Imry-Ma domains for a model with $\mathsf{O}(2)$ symmetry. Right: Energetics of domain formation. The region $L_{\rm d} < a$ is unphysical.

systems with continuous symmetries, such as the $\mathsf{O}(n)$ model, $V_{\rm dis} = -\sum_{\boldsymbol{r}} \boldsymbol{H}_{\boldsymbol{r}} \cdot \boldsymbol{S}_{\boldsymbol{r}}$. In charge density wave systems, if $\Psi(\boldsymbol{r})$ is the order parameter which encodes the local amplitude $|\Psi(\boldsymbol{r})|^2$ and phase $\arg\Psi(\boldsymbol{r})$ of the local density variation $\delta\varrho(\boldsymbol{r})$ relative to the homogeneous liquid, we may write

$$V_{\rm dis} = -\int d^2r\ \mathsf{Re}\left[H^*(\boldsymbol{r})\,\Psi(\boldsymbol{r})\right] \quad , \tag{3.86}$$

where $H(\boldsymbol{r}) \in \mathbb{C}$ is a complex number with a random amplitude and phase. Imry and Ma reasoned that such systems could try to lower their free energy by forming spatial domains in which the order parameter takes advantage of the local fluctuations in the random field. They presumed that such domains have a typical length scale $L_{\rm d}$, which is determined by the following energy minimization argument.

There are two contributions to the energy of a given domain: bulk and surface terms from the disorder. The bulk energy is given by $E_{\rm bulk} = -\Gamma(L_{\rm d}/a)^{d/2}$, where a is an ultraviolet cutoff, typically set by an atomic lattice spacing, and where $\Gamma = H_{\rm RMS} = \langle |H(\boldsymbol{r})|^2 \rangle$ is the root mean square amplitude of the random field. This is the Central Limit Theorem at work: if the phase of the CDW is locked over a patch of linear dimension $L_{\rm d}$, then adding $L_{\rm d}/a$ random fields gives us a contribution proportional to the square root of the number of such terms. The surface energy corresponds to the energy for creating a domain wall in the order parameter, which goes as

$$E_{\rm surf} \propto \begin{cases} J(L_{\rm d}/a)^{d-1} & \text{(discrete symmetry)} \\ J(L_{\rm d}/a)^{d-2} & \text{(continuous symmetry)} \end{cases} , \tag{3.87}$$

where J is the stiffness of the order parameter field arising from an energy density term $J|\boldsymbol{\nabla}\Psi|^2$. For the discrete case, the width of the domain wall may be taken to be a, in which case the surface energy is proportional to the surface area. For continuous symmetry, the domain wall is described by a continuous twist in the order parameter over a distance $L_{\rm d}$ perpendicular to the DW interface. If we take $\phi(x) = 2\pi x/L_{\rm d}$ and assume $|\Psi| \approx 1$ in the CDW, then

$$J\int\limits_0^{L_{\rm d}} dx\,(\partial_x\phi)^2 = \frac{4\pi^2 J}{L_{\rm d}} \quad , \tag{3.88}$$

which introduces a factor of $1/L_{\rm d}$ relative to the discrete case. Thus the free energy density per unit cell volume a^d is

$$f(L_{\rm d}) = \left(\frac{a}{L_{\rm d}}\right)^d \left(E_{\rm bulk} + E_{\rm surf}\right) \approx J\left(\frac{a}{L_{\rm d}}\right)^p - \Gamma\left(\frac{a}{L_{\rm d}}\right)^{d/2} \quad , \tag{3.89}$$

where $p = 1$ for discrete and $p = 2$ for continuous symmetry of the order parameter. Extremizing, we find that there is an extremum at

$$\frac{L_{\rm d}^*}{a} = \left(\frac{d_{\rm c}}{d}\cdot\frac{J}{\Gamma}\right)^{\frac{2}{d_{\rm c}-d}} \quad , \tag{3.90}$$

where $d_{\rm c} = 2p$ is $d_{\rm c} = 2$ (discrete) or $d_{\rm c} = 4$ (continuous). If $d < d_{\rm c}$, the extremum is a local maximum. For weak disorder, $\Gamma \ll J$, and thus $L_{\rm d} \gg a$. If $d > d_{\rm c}$, the free energy attains a local maximum at $L_{\rm d}^*$, but the sign of the exponent is reversed, and thus for weak disorder one has $L_{\rm d}^* \ll a$. Since $L_{\rm d}$ cannot become smaller than the UV cutoff scale a, the entire region $L_{\rm d} < a$ is unphysical, and the apparent instability where $f(L_{\rm d} \to 0) \to -\infty$ is avoided. The minimum value then occurs at $L_{\rm d}^* = \infty$, meaning that the LRO phase exists. The situation is summarized in Fig. 3.3.

Thus we conclude that a Wigner crystal phase in $d = 2$ with true LRO cannot exist in the presence of quenched disorder. Nevertheless, as we have seen, the length scale for Imry-Ma domains may be quite large, and it therefore makes good sense to speak of *local* crystalline order. Not included in the above analysis is the condensation energy of the ordered phase itself, which is dominated by local effects[14], and in assessing the stability of correlated liquid states, which we shall next discuss, it is generally sensible to compare to the energy density of the hypothetical pristine Wigner crystal.

[14] Although, as we have seen, the ground state energy density of triangular *versus* square Wigner crystals is slight and long-ranged features of the potential may possibly determine the specific crystallographic symmetry of the ordered state.

3.3 The Principal Sequence of Laughlin States

Two weeks before the publication of Laughlin's theory, Yoshioka, Halperin, and Lee[15] presented results of numerical studies which suggested that the ground state of the (Coulomb) interacting 2DEG at high magnetic fields does not possess solid-like order, and that the actual ground state energy lies somewhat below the corresponding Hartree-Fock CDW values[16]. They furthermore concluded,

> *We regard our data as supportive of the idea that the ground state is not crystalline, but a translationally invariant "liquid". We speculate that this liquid has commensuration energy at $\nu = \frac{1}{3}$ (and possibly at other simple rational values), and that for a large but finite system, the ground state at $\nu = \frac{1}{3}$ is threefold degenerate and separated by an energy gap from a variety of excited states. By going to a moving frame, it is then clear that at $\nu = \frac{1}{3}$ a Hall current will flow without dissipation, even in the presence of impurities. At ν close to $\frac{1}{3}$, we further suppose that the ground state, which is now highly degenerate, can be described as the $\nu = \frac{1}{3}$ ground state plus an additional small density of quasi "particles" or "holes". This leads naturally to a downward cusp in the energy as a function of ν. The Hall plateau at $\sigma_{xy} = e^2/3h$ can then be explained if the quasiparticles are localized by impurities and thus do not contribute to the Hall current, which is simply carried by the underlying $\nu = \frac{1}{3}$ state. Very recently, we have learned of a very original proposal by Laughlin of a wave function for a liquid state at $\nu = 1/p$, for p odd, which appears to have the requisite commensuration energy.*

One of the first good omens observed by Laughlin[17] was that his $\nu = \frac{1}{3}$ fluid ground state weighed in at a lower energy than did even the best CDW estimates. This sat well with those who viewed solid-like order with great uneasiness in light of earlier investigations of CDW pinning by disorder, which would be inconsistent with the observed finite (and indeed quantized) Hall conductivity at $\nu = \frac{1}{3}$. By comparing energies of the Laughlin fluid and the best correlated CDW states, a crude one-parameter phase diagram emerged, predicting a transition between correlated fluid and Wigner crystal phases at $\nu \approx \frac{1}{5}$ filling[18].

An argument by Allan MacDonald[19] concludes that the FQH state must be *incompressible* in the absence of disorder in order to be consistent with experiment. Recall that the isothermal compressibility of a thermodynamic system is defined to be

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = \frac{1}{n^2}\left(\frac{\partial n}{\partial \mu}\right)_T \quad . \tag{3.91}$$

[15] D. Yoshioka, B. I. Halperin, and P. A. Lee, *Phys. Rev. Lett.* **50**, 1219 (1983).

[16] The Maki-Zotos Wigner crystal describes the same state as the Hartree-Fock CDW and for $\nu < \frac{1}{2}$ has almost the same energy.

[17] R. B. Laughlin, *Phys. Rev. Lett.* **50**, 1395 (1983).

[18] As we have seen above (see Fig. 3.2, there is also experimental evidence for reentrant crystalline behavior.

[19] A. H. MacDonald, *Introduction to the Physics of the Quantum Hall Regime*, arXiv: cond-mat/9410047.

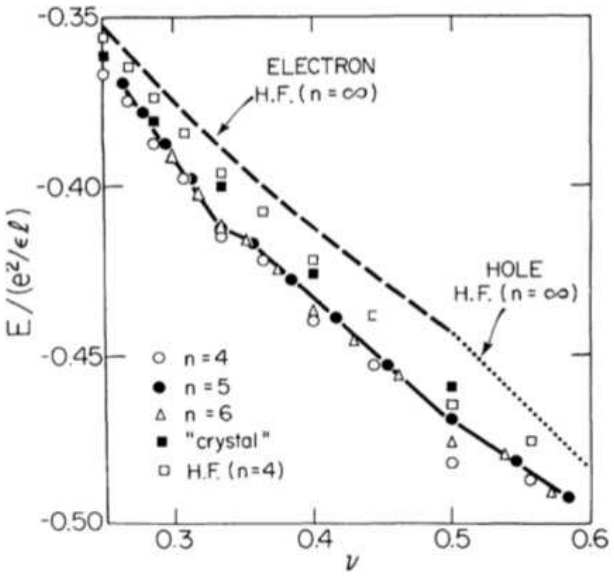


Figure 3.4: Energies for particle for the 2DEG as a function of ν, from D. Yoshioka, B. I. Halperin, P. A. and Lee, *Phys. Rev. Lett.* **50**, 1219 (1983). The dashed and dotted lines show electron and hole Hartree-Fock crystal energies for the infinite system. Open circles, closed circles, and triangles show exact diagonalization results for $N = 4$, 5, and 6 electrons. Energies for the crystalline state with $N = 4$ are shown with closed (exact diagonalization) and open (HF) squares. The solid line interpolating the $N = 5$ data is a guide to the eye.

It follows that when $\kappa = 0$, *i.e.* when the system is incompressible, the chemical potential μ is a discontinuous function of the density n. Consider now a quantum Hall droplet in which there is a current density $\boldsymbol{j}(\boldsymbol{r})$. The magnetization of the droplet is given by

$$\boldsymbol{M} = \frac{1}{2c}\int d^2r\, \boldsymbol{r}\times\boldsymbol{j}(\boldsymbol{r}) \qquad \Rightarrow \qquad \delta\boldsymbol{M} = \frac{1}{2c}\int d^2r\, \boldsymbol{r}\times\delta\boldsymbol{j}(\boldsymbol{r}) \quad . \tag{3.92}$$

Now let us imagine changing the chemical potential μ by an amount $\delta\mu$. If $\mu = \varepsilon_{\rm F}$ lies within a mobility gap, the only change in the current distribution can take place along the edge, which is some closed curve $\boldsymbol{R}(s)$. The parameterization is unimportant, but to be concrete we may take s to be the length along the curve. The differential change $\delta\boldsymbol{j}(\boldsymbol{r})$ in current density is then

$$\delta\boldsymbol{j}(\boldsymbol{r}) = \delta I\int d\boldsymbol{R}\,\delta(\boldsymbol{r}-\boldsymbol{R}) \quad , \tag{3.93}$$

where δI is the additional edge current. This entails the relation

$$\delta\boldsymbol{M} = \frac{\delta I}{2c}\oint \boldsymbol{R}\times d\boldsymbol{R} = \frac{A}{c}\,\delta I\,\hat{\boldsymbol{z}} \quad , \tag{3.94}$$

where A is the enclosed area. Thus,

$$\delta I = \frac{c}{A}\,\delta M = \frac{c}{A}\left(\frac{\partial M}{\partial\mu}\right)_B\delta\mu = \frac{c}{A}\left(\frac{\partial N}{\partial B}\right)_\mu\delta\mu \quad , \tag{3.95}$$

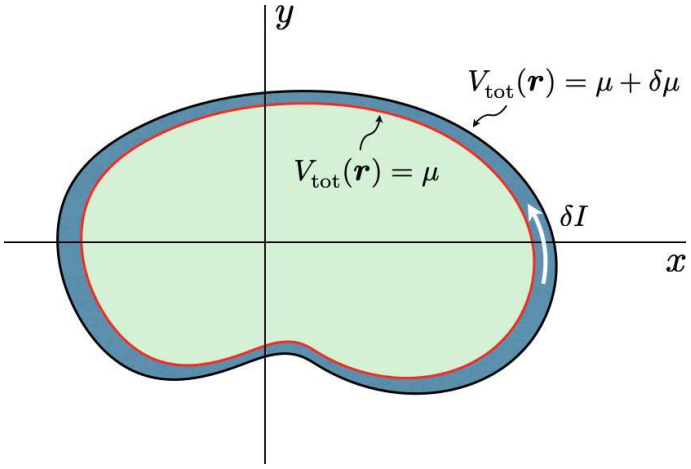


Figure 3.5: An expanding QH droplet, with $\delta \boldsymbol{j}(\boldsymbol{r}) = \delta I \int d\boldsymbol{R}\, \delta(\boldsymbol{r} - \boldsymbol{R})$. The single electron potential $V_{\rm tot}(\boldsymbol{r})$ is the sum of the random background and confining potentials plus Coulomb contributions from all the other electrons.

where we have invoked a Maxwell relation. This establishes the result

$$\frac{\delta I}{\delta \mu} = c \left(\frac{\partial n}{\partial B} \right)_\mu \quad . \tag{3.96}$$

Therefore, if $n(B)$ depends on B, *i.e.* if $(\partial n/\partial B)_\mu \neq 0$, and if μ lies within a spectral gap (or, more generally, a mobility gap) such that no internal currents are generated by an increase in chemical potential $\delta\mu$, then there must be gapless edge excitations.

As acknowledged by MacDonald, there are a couple of weak points to this argument. For example, it would seem that the Hall voltage $V_{\rm H}$ should have to be small in magnitude compared with $\hbar\omega_{\rm c}/e$, yet in experiments good quantization is observed even if the former is hundreds of times larger than the latter. In addition, it may not be that all the transport current flows at the edges of the system[20]. A more realistic approach to boundary conditions in the QHE was considered by Niu and Thouless[21].

3.3.1 Laughlin's excellent idea

Recall that all N-electron LLL states may be written in the form

$$\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = P(z_1, \ldots, z_N) \, \exp\left(- \frac{1}{4\ell^2} \sum_{j=1}^{N} |z_l|^2 \right) \quad , \tag{3.97}$$

[20] C. Wexler and D. J. Thouless, *Phys. Rev. B* **49**, 4815 (1994).
[21] Q. Niu and D. J. Thouless, *Phys. Rev. B* **35**, 2188 (1987).

where $P(Z)$ is a multinomial function which is odd under exchange, *i.e.*

$$P(Z_\sigma) = \mathsf{sgn}(\sigma) P(Z) \quad , \tag{3.98}$$

where $Z_\sigma = \{z_{\sigma(1)}, \ldots, z_{\sigma(N)}\}$, and where, restricting the single particle states to the angular momentum basis, with $m \in \{0, \ldots, N_\phi - 1\}$, the highest degree in each holomorphic coordinate z_j is $z_j^{N_\phi - 1}$. Furthermore, if Ψ is a state of definite total angular momentum $J = L_z$, then $F(Z)$ must be homogeneous, with $F(\lambda Z) = \lambda^{\mathsf{deg}(F)} F(Z)$ where $J = \mathsf{deg}(F)$. The filling fraction is then given by $\nu = N/N_\phi$. We've already encountered the example of the Vandermonde determinant,

$$V(Z) = \mathsf{det}(z_j^k) = \prod_{i>j} (z_i - z_j) \quad , \tag{3.99}$$

for which $J = \frac{1}{2} N(N-1)$ and $N_\phi = N$.

Laughlin (1983) proposed the sequence of FQHE states

$$\Psi_q(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \prod_{i>j} (z_i - z_j)^q \, \exp\left(- \frac{1}{4\ell^2} \sum_{j=1}^N |z_l|^2 \right) \quad , \tag{3.100}$$

i.e. $F_q(Z) = \big[V(Z)\big]^q$. Since $V(Z)$ is completely antisymmetric, we have

$$F_q(Z_\sigma) = \big(\mathsf{sgn}(\sigma)\big)^q F_q(Z) \quad , \tag{3.101}$$

hence Ψ_q corresponds to a fermionic wavefunction provided q is an odd integer. The total electronic angular momentum is $J = \frac{1}{2} qN(N-1)$ and the highest individual degree in any z_j is $N_\phi - 1 = q(N-1)$. Thus $\nu = N/N_\phi = q^{-1}$ in the thermodynamic limit. Note that for bosons, the Laughlin wavefunctions have m even.

3.3.2 Plasma analogy

The Laughlin states are a generalization of the Bijl-Jastrow pair product form, $\Psi_{\mathsf{BJ}} = \prod_{i<j} f(r_{ij})$. Consider, for example, the extended form,

$$\Psi_{\mathsf{GBJ}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \prod_i \exp\big\{ - \tfrac{1}{2} u(r_i) \big\} \prod_{i>j} \exp\big\{ - \tfrac{1}{2} v(r_{ij}) \big\} \prod_{i>j>k} \exp\big\{ - \tfrac{1}{2} w(r_{ij}, r_{jk}) \big\} \quad , \tag{3.102}$$

where $r_{ij} = |\boldsymbol{r}_i - \boldsymbol{r}_j|$. Then the N-particle probability density is

$$|\Psi_{\mathsf{GBJ}}|^2 = \exp\big\{ - \beta \Phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \big\} \tag{3.103}$$

where

$$\beta \Phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \sum_i u(r_i) + \sum_{i>j} v(r_{ij}) + \sum_{i>j>k} w(r_{ij}, r_{jk}) \quad . \tag{3.104}$$

Here we assume u, v, and w are all real. Thus the many-particle probability density is equivalent to the Boltzmann weight of a *classical* problem *in the same number of dimensions*, with one-body, two-body, three-body, *etc.* potentials.

For the Laughlin wavefunction, we have $|\Psi_q|^2 = \exp(-\beta\Phi)$, with $\beta = 1/q$ and

$$\Phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = -2q^2 \sum_{i>j} \ln|\boldsymbol{r}_i - \boldsymbol{r}_j| + \frac{q}{2\ell^2} \sum_i \boldsymbol{r}_i^2 \quad . \tag{3.105}$$

This is the classical *two-dimensional one-component plasma*, or 2DOCP, consisting of N point charges, each of strength $\theta = q\sqrt{2}$, interacting by the potential $v(\boldsymbol{r}) = \theta^2\phi(\boldsymbol{r}) = 2q^2\phi(\boldsymbol{r})$, with $\phi(\boldsymbol{r}) = -\ln r$, and subject to the background potential $u(\boldsymbol{r}) = q\boldsymbol{r}^2/2\ell^2$, all at temperature $k_{\rm B}T = q$. Note that $\nabla^2\phi(\boldsymbol{r}) = -2\pi\,\delta(\boldsymbol{r})$, hence $\nabla^2 u(\boldsymbol{r}) = 2q/\ell^2$, corresponding to the interaction of a charge $q\sqrt{2}$ with a uniform background of charge density $\rho = -1/\sqrt{2}\pi\ell^2$. To minimize the Coulomb energy, the N point charges form a disk of number density $n = 1/2\pi q\ell^2$, so that total charge neutrality holds, *i.e.* $nq\sqrt{2} + \rho = 0$. The radius R of this disk is then given by the condition $\pi R^2 n = N$, hence $R = \sqrt{2qN}\,\ell$.

3.3.3 The 2DOCP

Properties of the 2DOCP are discussed in a review article by J. M. Caillol *et al.*[22] To fix the problem precisely, consider a classical system of particles each of charge e, interacting via a potential $v(r) = -e^2\ln(r/d)$, where d is a length scale. Clearly d is irrelevant as it enters the energy additively and thus sets the location of the zero of energy. Consider N such particles in a disk of radius R. The mean particle number density is thus $n = N/\pi R^2$; then $a \equiv (\pi n)^{-1/2}$ is called the *ion disk radius*. Let the disk be filled with a uniform neutralizing background of charge density $(-en)$. Taking into account particle-particle, particle-background, and background-background interactions, the energy is then

$$H(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = -e^2 \sum_{j<k}^{N} \ln\left(\frac{|\boldsymbol{r}_j - \boldsymbol{r}_k|}{d}\right) + \tfrac{1}{2}Ne^2 \sum_{i=1}^{N} \left(\frac{r_i}{R}\right)^2 + \tfrac{1}{2}N^2e^2\left[\ln\left(\frac{R}{d}\right) - \frac{3}{4}\right] \tag{3.106}$$

The partition function is

$$Z = \int^R d^2r_1 \cdots \int^R d^2r_N\, e^{-\beta H(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)} \equiv e^{-N\beta f} \quad , \tag{3.107}$$

where $f = -N^{-1}k_{\rm B}T\ln Z$ is the free energy per particle. Defining $\boldsymbol{x}_i \equiv \boldsymbol{r}_i/R$,

$$\beta f = (1 - \tfrac{1}{4}\,\Gamma)\ln(\pi n) - \tfrac{1}{2}\,\Gamma\ln d - \tfrac{3}{8}\,N\,\Gamma + \tfrac{1}{4}\,\Gamma\ln N - \frac{1}{N}\ln W(N, \Gamma) \quad , \tag{3.108}$$

[22] J. M. Caillol, D. Levesque, J. J. Weis, and J. P. Hansen, *J. Stat. Phys.* **28**, 325 (1982).

where $\Gamma \equiv \beta e^2$ is the dimensionless *plasma parameter* and

$$W(N,\Gamma) = \int^1 d^2x_1 \cdots \int^1 d^2x_N \prod_{j<k}^{N} |\boldsymbol{x}_j - \boldsymbol{x}_k|^\Gamma \prod_{i=1}^{N} e^{-N\Gamma \boldsymbol{x}_i^2/2} \quad . \tag{3.109}$$

From the thermodynamic relation

$$df\big|_N = -s\,dT + pn^{-2}dn \quad , \tag{3.110}$$

we have the equation of state

$$p = (1 - \tfrac{1}{4}\,\Gamma)\,nk_{\rm B}T \quad . \tag{3.111}$$

The isothermal compressibility is then $\kappa_T = n^{-1}(\partial n/\partial p)_T = 1/p$, and we define the dimensionless isothermal compressibility $\chi_T \equiv nk_{\rm B}T\,\kappa_T = (1 - \frac{1}{4}\,\Gamma)^{-1}$.

The equilibrium properties of the plasma are dependent solely on Γ. When Γ is small, the plasma is said to be weakly coupled. For $\Gamma > \Gamma_{\rm c} \approx 140$, the 2DOCP crystallizes. Much of the physics of the 2DOCP is reflected in the behavior of the pair distribution function,

$$n\,g(\boldsymbol{r}) = \frac{1}{N}\left\langle \sum_{i\neq j}^{N} \delta(\boldsymbol{r} + \boldsymbol{r}_j - \boldsymbol{r}_i) \right\rangle \tag{3.112}$$

and the associated pair correlation function $h(\boldsymbol{r}) = g(\boldsymbol{r}) - 1$. The static structure factor, for example, is given by

$$\hat{s}(\boldsymbol{k}) = 1 + n\int d^2r\; h(\boldsymbol{r})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} = 1 + n\,\hat{h}(\boldsymbol{k}) \quad . \tag{3.113}$$

The Fourier transform of the *direct correlation function*, $\hat{c}(\boldsymbol{k})$, is defined by

$$\hat{c}(\boldsymbol{k}) \equiv \frac{\hat{h}(\boldsymbol{k})}{1 + n\,\hat{h}(\boldsymbol{k})} \qquad \Longleftrightarrow \qquad \hat{h}(\boldsymbol{k}) = \frac{\hat{c}(\boldsymbol{k})}{1 - n\,\hat{c}(\boldsymbol{k})} \quad , \tag{3.114}$$

which is equivalent to the relation

$$h(\boldsymbol{r}) = c(\boldsymbol{r}) + n\int d^3r'\; h(\boldsymbol{r} - \boldsymbol{r}')\, c(\boldsymbol{r}') \quad . \tag{3.115}$$

This is known as the Ornstein-Zernike equation. Physically it says that the correlation between a particle at position $\mathbf{0}$ and a particle at position $\boldsymbol{r}$ can be written as a sum of the direct correlator $c(\boldsymbol{r})$ plus an term arising from the indirect effect of the particle at $\mathbf{0}$ with a third particle at $\boldsymbol{r}'$ which affects that at $\boldsymbol{r}$ both directly and indirectly.

The asymptotic long wavelength behavior of $\hat{c}(\boldsymbol{k})$ is believed to correspond to a weak coupling limit, in which case

$$\hat{c}(\boldsymbol{k}) \longrightarrow -\beta\hat{v}(\boldsymbol{k}) = -Q^2/nk^2 \quad , \tag{3.116}$$

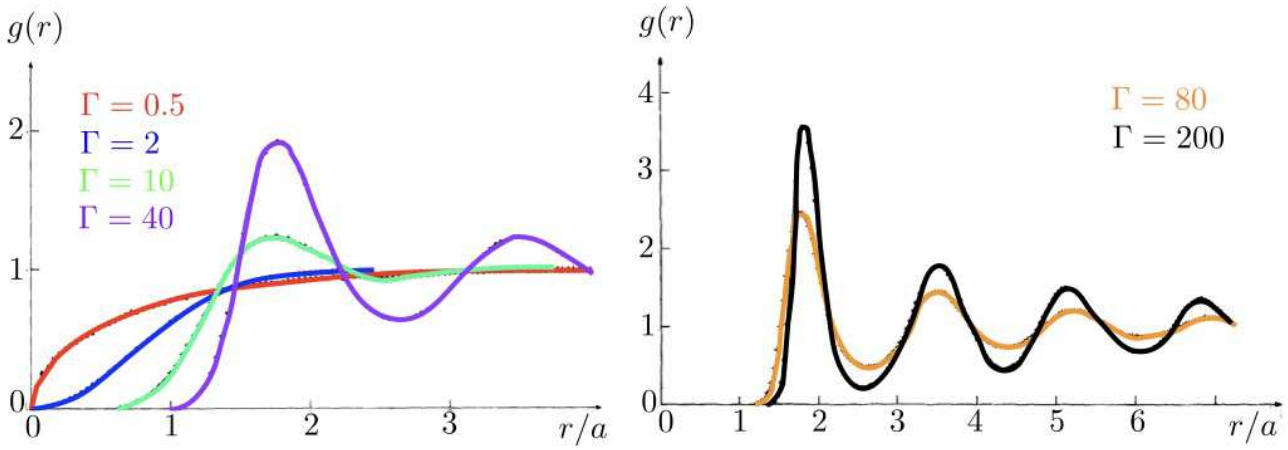


Figure 3.6: Pair distribution function $g(r)$ for the 2DOCP. Monte Carlo results from Figs. 1 and 2 of Caillol *et al.* (1982).

where $Q = (2\pi n\beta e^2)^{1/2} = (2\pi n\Gamma)^{1/2}$ is the Debye wavevector. Separating out this singular term, one writes $\hat{c}(\boldsymbol{k}) = -Q^2/nk^2 + \hat{c}^{\rm R}(\boldsymbol{k})$ where $\hat{c}^{\rm R}(\boldsymbol{k}) = \hat{c}^{\rm R}(0) + \mathcal{O}(\boldsymbol{k}^2)$. The regular part $\hat{c}^{\rm R}(\boldsymbol{k})$ is related to the dimensionless isothermal compressibility:

$$\lim_{\boldsymbol{k}\to 0} \hat{c}^{\rm R}(\boldsymbol{k}) = \hat{c}^{\rm R}(0) = n^{-1}(1 - \chi_T^{-1}) \quad . \tag{3.117}$$

The above results entail the expansion

$$n\,\hat{h}(\boldsymbol{k}) = -1 + \frac{k^2}{Q^2} + \frac{k^4}{\chi_T Q^4} + \mathcal{O}(k^6) \quad . \tag{3.118}$$

In real space, then, we have the following sum rules on moments of the pair correlation function:

$$n\int d^2r\; h(\boldsymbol{r}) = n\,\hat{h}(\boldsymbol{q})\big|_{\boldsymbol{q}=0} = -1 \quad , \tag{3.119}$$

known as the *charge neutrality* sum rule,

$$n\int d^2r\; r^2\, h(\boldsymbol{r}) = -n\,\boldsymbol{\nabla}_{\boldsymbol{q}}^2\,\hat{h}(\boldsymbol{q})\big|_{\boldsymbol{q}=0} = -\frac{4}{Q^2} \quad , \tag{3.120}$$

known as the *perfect screening* sum rule, and

$$n\int d^2r\; r^4\, h(\boldsymbol{r}) = n\,(\boldsymbol{\nabla}_{\boldsymbol{q}}^2)^2\,\hat{h}(\boldsymbol{q})\big|_{\boldsymbol{q}=0} = -\frac{64}{\chi_T Q^4} \quad , \tag{3.121}$$

known as the *compressibility* sum rule.

From Monte Carlo studies, we know that the 2DOCP crystallizes at $\Gamma \simeq 140$. At this point, the structure factor $s(\boldsymbol{k})$ exhibits Bragg peaks. Since the Coulomb potential is long-ranged, the system can evade the usual Hohenberg-Mermin-Wagner restrictions which forbid broken continuous translational symmetry at finite temperature in $d \leq 2$ dimensions.

3.3.4 Laughlin *vs.* Wigner crystal

The correspondence between the Laughlin wavefunction $|\Psi_q|^2 = \exp(-\beta H)$ and the 2DOCP Hamiltonian H is then:

$$\beta = \frac{1}{q} \quad , \quad e^2 = 2q^2 \qquad \Gamma = 2q \quad , \quad n = \frac{1}{2\pi q \ell^2} \quad . \tag{3.122}$$

The sum rules provide information on the long-wavelength behavior of the structure factor in the Laughlin states, *viz.*

$$\hat{s}(\boldsymbol{k}) = \tfrac{1}{2}\, k^2\ell^2 + \tfrac{1}{8}\,(q-1)\, k^4\ell^4 + \mathcal{O}(k^6) \quad . \tag{3.123}$$

Crystallization of the 2DOCP (into a triangular structure) at $\Gamma = 140$ means that the Laughlin wavefunction has triangular crystalline order for $q > 70$. However, recall that the Laughlin state is an *Ansatz* wavefunction. It isn't even a proper variational state, since the only free parameter m is discrete and is fixed by the filling, with $\nu = q^{-1}$.

The actual 2DEG in the LLL crystallizes well before the filling fraction gets as low as $\frac{1}{70}$. Upon taking the thermodynamic limit and properly including the effects of the uniform neutralizing background, the energy per particle is given by

$$u = \frac{U}{N} = \frac{n}{2}\int d^2r\; v(\boldsymbol{r})\,\big[g(\boldsymbol{r}) - 1\big] \quad , \tag{3.124}$$

where $v(r) = e^2/\epsilon r$. The pair distribution function in the Laughlin states has been evaluated using the so-called hypernetted chain approximation (Laughlin, 1983) and by Monte Carlo methods[23]. At $\nu = \frac{1}{3}$, these calculations yield energies $u_{\rm L}^{\rm HNC}(q=3) = -0.4156 \pm 0.0012$ and $u_{\rm L}^{\rm MC}(q=3) = -0.410 \pm 0.001$, respectively (units of $e^2/\epsilon\ell$). Exact diagonalization studies by Haldane and Rezayi[24] extrapolated finite size results on the sphere for $N \leq 7$ to $N = \infty$ and obtained $u_{\rm L}^{\rm ED}(q=3) = -0.415 \pm 0.005$. In contrast, the energy for the Wigner crystal, as computed by Yoshioka, Halperin, and Lee, or by Maki and Zotos, is about $u^{\rm WC}(\nu = \frac{1}{3}) \approx -0.38$, which is much higher. In order to give the Wigner crystal a fighting chance, Lam and Girvin[25] investigated the correlated Wigner crystal wavefunction,

$$\Psi_{\rm CWC}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \exp\left(\frac{1}{2}\sum_{i,j}(z_i - R_i)\, B_{ij}\, (z_j - R_j)\right)\prod_l \varphi_{\boldsymbol{R}_l}(\boldsymbol{r}_l) \quad , \tag{3.125}$$

where the prefactor is a holomorphic function in $\{z_1, \ldots, z_N\}$ and $B_{ij} = B(\boldsymbol{R}_i - \boldsymbol{R}_j)$ is a correlating matrix. As the effects of antisymmetrization below $\nu = 0.4$ were found by Maki and

[23] See D. Levesque, J. J. Weis, and A. H. MacDonald, *Phys. Rev. B* **30**, 1056 (1984) and R. Morf and B. I. Halperin, *Phys. Rev. B* 32, 2221 (1986).

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

[25] P. K. Lam and S. M. Girvin, *Phys. Rev. 30*, 473 (1984).

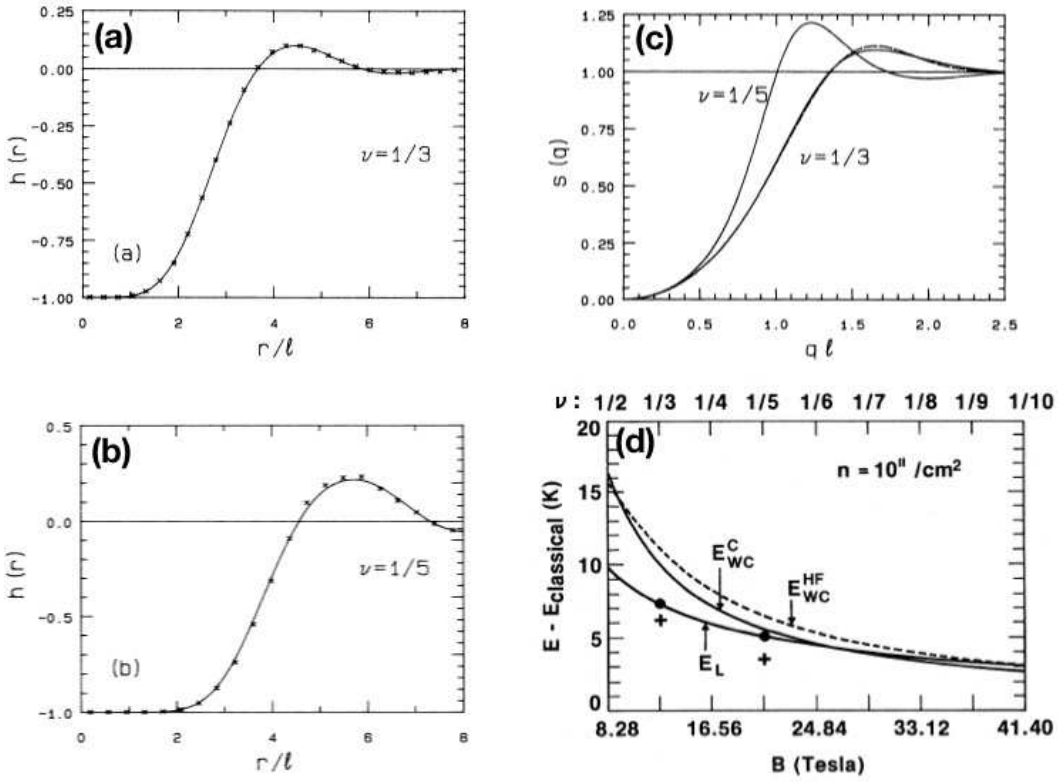


Figure 3.7: Pair correlation function (a,b) and structure factor (c) for $q = 3$ and $q = 5$ Laughlin states. From S. M. Girvin, A. M. MacDonald, and P. M. Platzman, *Phys. Rev. B* **33**, 2481 (1986). Panel (d) shows a comparison of the (interpolated Laughlin state energy with that of the correlated Wigner crystal, indicating a transition to the WC state at $\nu^{-1} \approx 6.5$. From P. K. Lam and S. M. Girvin, *Phys. Rev. 30*, 473 (1984).

Zotos to be insignificant, no antisymmetrization was imposed. The matrix B_{ij} was then determined variationally by minimizing the Coulomb energy in this state. Within the harmonic approximation, the lattice Fourier transform $\hat{B}(\boldsymbol{k}) = \sum_{\boldsymbol{R}} B(\boldsymbol{R})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{R}}$ is given by

$$\hat{B}(\boldsymbol{k}) = \frac{\omega_{\mathsf{L}}(\boldsymbol{k}) - \omega_{\mathsf{T}}(\boldsymbol{k})}{\omega_{\mathsf{L}}(\boldsymbol{k}) + \omega_{\mathsf{T}}(\boldsymbol{k})} \quad , \tag{3.126}$$

where $\omega_{\mathsf{L,T}}(\boldsymbol{k})$ are the longitudinal and transverse magnetophonon frequencies. Such an optimized WC state significantly lowers the correlation energy, to $u_{\mathsf{CWC}}(\nu = \frac{1}{3}) = -0.3948 \pm 0.0005$, which is still about 2.7% higher than that of the $q = 3$ Laughlin state. The Laughlin state is remarkably robust in terms of its Coulomb energy with respect to variational tweaking, and the extrapolated differences between the Coulomb energy in the Laughlin state and that from exact diagonalization differ by about 0.05%[26].

[26]See M. Fremling, J. Fulsebakke, N. Moran, and J. K. Slingerland, *Phys. Rev. B*, **93**, 235149 (2016).

3.3.5 Haldane pseudopotentials

It was realized by Haldane[27] and by Trugman and Kivelson[28] that the Laughlin state Ψ_q is the *exact* ground state for sufficiently short-ranged interaction potentials[29]. Whenever the interaction potential $v(\boldsymbol{r})$ is central, *i.e.* a function of $r = |\boldsymbol{r}|$ alone, its Fourier transform,

$$\hat{v}(\boldsymbol{k}) = \int d^2r\, v(\boldsymbol{r})\, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} = 2\pi \int_0^\infty dr\, r\, v(r) J_0(kr) \quad , \tag{3.127}$$

may be expanded as a power series in k^2, *i.e.* $\hat{v}(\boldsymbol{k}) = \sum_{j=0}^\infty A_j\, (-k^2\ell^2)^j$. Thus in real space, we may write $v(\boldsymbol{r})$ as an expansion in powers of the Laplacian acting on a delta-function:

$$v(\boldsymbol{r}) = \sum_{j=0}^\infty A_j\, (\ell^2\, \nabla^2)^j\, \delta(\boldsymbol{r}) \quad . \tag{3.128}$$

For a totally antisymmetric wavefunction $\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)$, the $j = 0$ term will not contribute to the total energy.

The interaction energy per particle is given by[30]

$$\frac{E_{\rm int}}{N} = \langle \Psi | H_{\rm int} | \Psi \rangle = \tfrac{1}{2} n \int d^2r\, v(\boldsymbol{r})\, g(\boldsymbol{r}) = \tfrac{1}{2} n \sum_{j=0}^\infty A_j\, (\ell^2 \nabla^2)^j g(\boldsymbol{r}) \Big|_{\boldsymbol{r}=0} \quad , \tag{3.129}$$

where $g(\boldsymbol{r})$ is the pair distribution function. In the Laughlin state Ψ_q, the $r \to 0$ behavior of $g(\boldsymbol{r})$ is given by

$$g(\boldsymbol{r}) = c_q\, (r/\ell)^{2q} + c_{q+1}\, (r/\ell)^{2(q+1)} + \ldots \quad , \tag{3.130}$$

where the $\{c_j\}$ are constant coefficients. Thus all terms in the potential with coefficients A_j with $j < q$ will make no contribution to $E_{\rm int}$. Suppose now that $A_{1,2} > 0$ but $A_3 = A_4 = \cdots = 0$. If a wavefunction $\Psi = P(Z) \prod_j e^{-|z_j|^2/4\ell^2}$ is to satisfy $\langle \Psi | H_{\rm int} | \Psi \rangle = 0$, it is clear that it must vanish at least as fast as $(z_i - z_j)^3$ as the pair separation tends to zero, and hence $P(Z)$ must contain at least 3 factors of $V(Z)$, where $V(Z)$ is the Vandermonde determinant. On the other hand, for homogeneous states, the filling factor is given by

$$\nu = \frac{N(N-1)}{2J} \quad , \tag{3.131}$$

[27] F. D. M. Haldane, *Phys. Rev. Lett.* **51**, 605 (1983).

[28] S. A. Trugman and S. Kivelson, *Phys. Rev. B* **31**, 5280 (1985).

[29] On the sphere, the Laughlin state is nondegenerate. On the torus, the degeneracy is q, and on a Riemann surface of genus g, the degeneracy is q^g. See X. G. Wen and Q. Niu, *Phys. Rev. B* **41**, 9377 (1990). On the plane, with a confining potential, there will be a continuum of gapless edge excitations.

[30] This energy includes only particle-particle interactions and does not include the effects of any neutralizing background.

where J is the total angular momentum. Since $\deg(V) = \frac{1}{2}N(N-1)$, we conclude that the holomorphic part of Ψ is $P(Z) = C\big[V(Z)\big]^q$, where C is a constant. Thus the Laughlin state with $q = 3$ is the *only* homogeneous state at $\nu = \frac{1}{3}$ which has zero energy. Similarly, when $A_{1,2,3,4} > 0$ and $A_5 = A_6 = \cdots = 0$, the $q = 5$ Laughlin state is the sole $\nu = \frac{1}{5}$ state which lies at zero energy[31].

There is a convenient parameterization (Haldane 1983) of the pair interaction in terms of relative coordinate 'pseudopotentials'. Writing

$$v(\boldsymbol{r}_i - \boldsymbol{r}_j) = \int \frac{d^2k}{(2\pi)^2}\, e^{i\boldsymbol{k}\cdot(\boldsymbol{r}_i - \boldsymbol{r}_j)} \quad , \tag{3.132}$$

and invoking the separation into cyclotron and guiding center ladder operators[32] $z = \sqrt{2}\,\ell(a + b^\dagger)$, we have that the LLL-projected interaction is

$$\begin{aligned}\Pi_0\, v(\boldsymbol{r}_i - \boldsymbol{r}_j)\,\Pi_0 &= \int \frac{d^2k}{(2\pi)^2}\,\hat{v}(\boldsymbol{k})\,\exp\left[\frac{i\ell k}{\sqrt{2}}(b_i - b_j)\right]\exp\left[\frac{i\ell\bar{k}}{\sqrt{2}}(b_i^\dagger - b_j^\dagger)\right] \\ &= \sum_{n=0}^{\infty}\frac{1}{(n!)^2}\int\frac{d^2k}{(2\pi)^2}\,\hat{v}(\boldsymbol{k})\,(-k^2\ell^2)^n\,(J_{ij}^-)^n\,(J_{ij}^+)^n \quad .\end{aligned} \tag{3.133}$$

The operators $J_{ij}^\pm$ raise and lower the *relative angular momentum* of the pair (ij):

$$J_{ij}^+ = \tfrac{1}{\sqrt{2}}(b_i^\dagger - b_j^\dagger) \qquad , \qquad J_{ij}^- = \tfrac{1}{\sqrt{2}}(b_i - b_j) \quad , \tag{3.134}$$

with

$$J_{ij} = \Pi_0\Big(\tfrac{1}{2}(\boldsymbol{r}_i - \boldsymbol{r}_j)\times(\boldsymbol{p}_i - \boldsymbol{p}_j)\cdot\hat{\boldsymbol{z}}\Big)\Pi_0 = J_{ij}^+\,J_{ij}^- \quad . \tag{3.135}$$

The relative angular momentum raising and lowering operators satisfy the algebra

$$\big[J_{ij}^-, J_{ij}^+\big] = \tfrac{1}{2}\big(\delta_{ik} + \delta_{jl} - \delta_{il} - \delta_{jk}\big) \quad . \tag{3.136}$$

If we define the projector $\mathrm{P}_s(ij)$ to be the projector of the pair (ij) onto the relative angular momentum $J_{ij} = s$ subspace, then

$$\begin{aligned}\Pi_0\, v(\boldsymbol{r}_i - \boldsymbol{r}_j)\,\Pi_0 &= \sum_{s=0}^{\infty} V_s\,\mathrm{P}_s(ij) \\ V_s &= \int\frac{d^2k}{(2\pi)^2}\,\hat{v}(\boldsymbol{k})\,L_s(\boldsymbol{k}^2\ell^2)\,e^{-\boldsymbol{k}^2\ell^2} \quad ,\end{aligned} \tag{3.137}$$

where $L_s(x)$ is the Laguerre polynomial. This provides us with another way to expand the real space potential, *i.e.* in terms of the pseudopotential amplitudes $\{V_s\}$:

$$v(\boldsymbol{r}) = 4\pi\ell^2\sum_{s=0}^{\infty} V_s\,L_s(-\ell^2\nabla^2)\,\delta(\boldsymbol{r}) \quad . \tag{3.138}$$

[31] Other than center-of-mass degeneracies on Riemann surfaces of genus $g > 0$.

[32] Recall we have been working in the symmetric gauge since §1.3.7.

which may be compared with Eqn. 3.128. One then finds

$$A_j = \frac{4\pi\ell^2}{j!}\sum_{s=j}^{\infty}\binom{s}{s-j}V_s \quad . \tag{3.139}$$

If $V_s = 0$ for all $s > s_0$, then the power series expansion of Eqn. 3.128 is simply a rearrangement of the pseudopotential expansion, up to the same leading order in powers of ∇^2. Again, only odd pseudopotentials contribute to the energy for a fermionic system.

For the Coulomb interaction $v(\boldsymbol{r}) = e^2/\epsilon r$, one finds

$$V_s^{\text{COUL}} = \frac{1}{4^s}\binom{2s}{s}\cdot\frac{\sqrt{\pi}e^2}{2\epsilon\ell} \quad . \tag{3.140}$$

As $s \to \infty$, we have $V_s \simeq e^2/2\sqrt{s}\epsilon\ell \propto s^{-1/2}$, corresponding to the fact that the single particle state with angular momentum s encloses an area $\pi r^2 = 2\pi s\ell^2$. Thus $V_1^{\text{COUL}} = \sqrt{\pi}e^2/4\epsilon\ell$ and $V_3^{\text{COUL}} = \frac{5}{8}\,V_1^{\text{COUL}}$, $V_5^{\text{COUL}} = \frac{63}{238}\,V_1^{\text{COUL}}$, *etc.* If we define

$$V_s(\lambda) = (1-\lambda)V_1^{\text{COUL}}\,\delta_{s,1} + \lambda\,V_s^{\text{COUL}} \quad , \tag{3.141}$$

which interpolates between the truncated pseudopotential $V_s(0) = V_1^{\text{COUL}}\,\delta_{s,1}$ at $\lambda = 0$ and the full Coulomb $V_s(1) = V_s^{\text{COUL}}$ at $\lambda = 1$. One can then ask whether the gap at $\lambda = 0$ remains finite for all $\lambda \in [0,1]$, in which case no phase boundaries are crossed as one evolves from the truncated pure V_1 model to the full Coulomb interaction. Alternatively, Haldane[33] considered the potential given by

$$V_s = \big(V_1 - V_1^{\text{COUL}}\big)\delta_{s,1} + V_s^{\text{COUL}} \tag{3.142}$$

as a function of V_1, which takes the Coulomb interaction and replaces the $s = 1$ pseudopotential component V_1^{COUL} with the parameter V_1. Results for $N = 6$ particles in a spherical geometry are shown in Fig. 3.8. As the V_1 pseudopotential is decreased from its Coulomb value of $V_1^{\text{COUL}} = 0.4781$ on the sphere[34], the bulk gap is found to collapse at $V_3 \approx 0.37$, heralding a second order phase transition to a compressible phase.

To understand better the spectrum of relative angular momentum in the Laughlin states, define the complex center-of-mass and relative coordinates for a select pair $(i = 1, j = 2)$ as $W \equiv \frac{1}{2}(z_1 + z_2)$ and $w \equiv z_2 - z_1$, so that $z_{1,2} = W \mp \frac{1}{2}w$. Then

$$V(z_1,\ldots,z_N) = w\prod_{j=3}^{N}\Big[(W-z_j)^2 - \tfrac{1}{4}w^2\Big]\,V(z_3,\ldots,z_N) \quad . \tag{3.143}$$

The spectrum of relative angular momentum states for the pair $(1,2)$ can be gleaned by identifying terms homogeneous in the relative coordinate w, *i.e.* terms proportional to w^l for some l.

[33] F. D. M. Haldane in *The Quantum Hall Effect*, R. E. Prange and S. M. Girvin, eds. (Springer, 1987).

[34] On the sphere, the Coulomb interaction is proportional to the inverse chord length, yielding $V_1^{\text{COUL}} = 0.4781$ and $V_3^{\text{COUL}} = 0.3092$, rather than $V_1^{\text{COUL}} = 0.4431$ and $V_3^{\text{COUL}} = 0.2769$ as obtained on the plane.

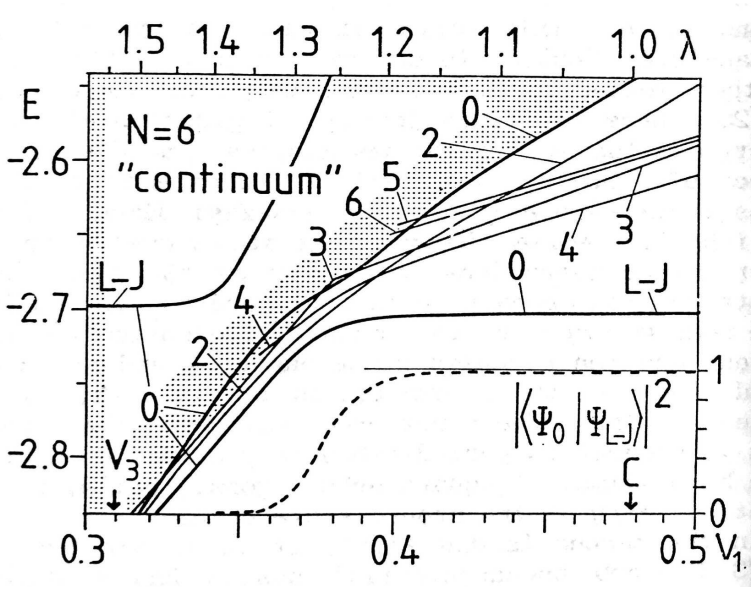


Figure 3.8: Effect of varying the V_1 pseudopotential on low-lying energy states. Excited energy levels are labeled by their total angular momentum L on the sphere ($N = 6$). The Laughlin-Jastrow (L-J) state has $L = 0$. The overlap of the ground state with the Laughlin state is also shown. Arrows on the V_1 axis indicate values of V_1^{COUL} and V_3^{COUL} (units of $e^2/\epsilon\ell$). For $V_1 > 0.37$ the system is gapped and incompressible. Below this value the system is gapless and hence compressible. From Haldane (1987).

Clearly we have contributions from $l \in \{1, 3, 5, \ldots, 2N-3\}$. Note that only odd l terms enter the spectrum. When we raise $V(Z)$ to the power q, we obtain

$$l \in \big\{q, q+2, q+4\,,\ \ldots\ ,\ 2q(N-2)+1\big\} \quad . \tag{3.144}$$

Thus, in the Laughlin state Ψ_q, the spectrum of relative coordinate angular momenta is all odd integers l starting at $l_{\min} = q$, and terminating at the cutoff $l_{\max} = 2q(N-1)+1$.

Given an arbitrary many-body state, the pair distribution function $g(r)$ may be expanded in powers of r^2 for small r, *viz.*

$$g(r) = \sum_{p=0}^{\infty} c_p\,(r^2/\ell^2)^p \quad . \tag{3.145}$$

If the holomorphic factor $P(Z)$ contains q factors of the Vandermonde determinant $V(Z)$, then $c_p = 0$ for $p = 0, \ldots, q-1$. Exact diagonalization studies by Yoshioka[35] for small fermionic ($N \leq 8$) and bosonic ($N \leq 7$) systems in a toroidal geometry examined the behavior of the coefficients c_p as a function of filling fraction. Results are shown in Fig. 3.9. One sees that for fermionic states with $\nu < \frac{1}{3}$, the coefficients $c_{1,2}$ are both exceedingly small, suggesting that the cube of the Vandermonde determinant approximately divides the holomorphic part

[35]D. Yoshioka, *Phys. Rev. B* **29**, 6822 (1984).

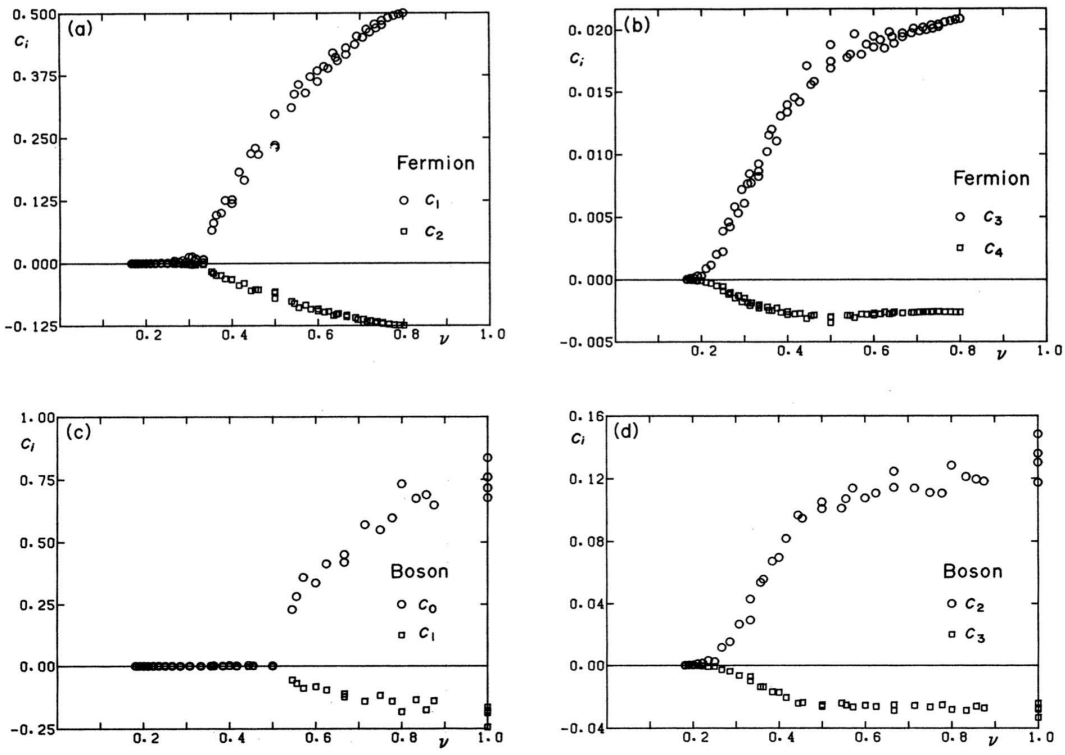


Figure 3.9: Pair distribution function coefficients in exact finite N ground states *versus* filling fraction. (a) Coefficients c_1 (circles) and c_2 (squares) for fermionic states. (b) Coefficients c_3 (circles) and c_4 (squares) for fermionic states. (c) Coefficients c_0 (circles) and c_1 (squares) for bosonic states. (d) Coefficients c_2 (circles) and c_3 (squares) for bosonic states. From D. Yoshioka, *Phys. Rev. B* **29**, 6822 (1984).

$P(Z)$ of the exact wavefunction. Similar results are found *vis-a-vis* $c_{3,4}$ when $\nu < \frac{1}{5}$. Yoshioka's numerics establish that for Coulomb systems with $\nu \leq q^{-1}$, for both bosons and fermions, the holomorphic part $P(Z)$ of the ground state wavefunction is almost perfectly divided by the $q^{\rm th}$ power of the Vandermonde determinant.

3.3.6 Quasiparticles

Laughlin also proposed wavefunctions for localized charged excitations, called *quasiholes* and *quasielectrons*. The quasihole wavefunctions are the easiest to understand and are of the form

$$\Psi_q^{\rm QH}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N; \boldsymbol{\xi}) = \prod_{l=1}^{N}(z_l - \xi) \prod_{j>k}^{N}(z_j - z_k)^q \prod_{i=1}^{N} \exp(-z_i \bar{z}_i / 4\ell^2) \quad , \tag{3.146}$$

where ξ is the complexified quasihole position $\xi = \xi_x + i\xi_y$. Laughlin argues that such a state should be gauge-equivalent to an eigenstate of a many-electron system in which Ψ_q is a ground

state, since it results from adiabatic flux threading by $\phi_0 = hc/e$ *parallel to the applied field* at the point $\boldsymbol{r} = \boldsymbol{\xi}$. The associated plasma Hamiltonian is then

$$\Phi_q(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N; \boldsymbol{\xi}) = -m \ln |\Psi_q|^2 = -2q^2 \sum_{j>k} \ln |\boldsymbol{r}_j - \boldsymbol{r}_k| + \frac{q}{2\ell^2} \sum_i \boldsymbol{r}_i^2 - 2q \sum_l \ln |\boldsymbol{r}_l - \boldsymbol{\xi}| \quad . \tag{3.147}$$

The last term corresponds to a charge $\theta^* = \sqrt{2}$ object at $\boldsymbol{\xi}$ interacting via the logarithmic potential $v(\boldsymbol{r}) = -\theta\theta^* \ln |\boldsymbol{r}_i - \boldsymbol{\xi}|$ with charge $e^* = q\sqrt{2}$ particles at each $\boldsymbol{r}_i$. Since any charged impurities are completely screened in the plasma phase of the 2DOCP, there will be a deficit of $\theta^*/\theta = q^{-1}$ particles localized about the position $\boldsymbol{\xi}$ on the scale of the Debye length, which is $\lambda_{\mathsf{D}} = 1/Q = a/\sqrt{2\Gamma}$ where $a = (\pi n)^{-1/2} = \sqrt{2q}\,\ell$ is the ion disk radius of the 2DOCP and $\Gamma = 2q$ is the plasma parameter. Thus $\lambda_{\mathsf{D}} = \ell/\sqrt{2}$ is the screening length.

The quasielectron wavefunctions are somewhat more complicated. Adiabatically inserting flux ϕ_0 parallel to the applied field has the effect of $|\, m \,\rangle \to |\, m+1 \,\rangle$ on the single particle angular momentum basis states with origin at $\boldsymbol{\xi}$. If we adiabatically insert flux ϕ_0 *antiparallel* to the applied field, it stands to reason that $|\, m \,\rangle \to |\, m-1 \,\rangle$, in which case whither $|\, m = 0 \,\rangle$? Laughlin's proposed quasielectron wavefunction is given by

$$\Psi_q^{\mathsf{QE}}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N; \boldsymbol{\eta}) = \prod_{i=1}^N \exp(-z_i \bar{z}_i / 4\ell^2) \prod_{l=1}^N \left(2\ell^2 \frac{\partial}{\partial z_l} - \bar{\eta} \right) \prod_{j>k}^N (z_j - z_k)^q \quad , \tag{3.148}$$

where $\bar{\eta} = \eta_x - i\eta_y$. This form is inspired by the Girvin-Jach substitution $\bar{z} \to 2\ell^2 \partial$, and results in a localized defect of increased number density q^{-1} within λ_D of the quasielectron at $\boldsymbol{\eta}$. Note that dividing Ψ_q by the product $\prod_j (z_j - \eta)$ results in a nonanalyticity and the resulting state is no longer in the LLL. Morf and Halperin (1986) calculated the quasihole and quasielectron energies at $\nu = \frac{1}{3}$ via Monte Carlo, obtaining

$$\tilde{\varepsilon}_{\mathrm{MH}}^{\mathsf{QH}} = (0.0268 \pm 0.0033) \frac{e^2}{\epsilon\ell} \qquad , \qquad \tilde{\varepsilon}_{\mathrm{MH}}^{\mathsf{QP}} = (0.073 \pm 0.008) \frac{e^2}{\epsilon\ell} \quad . \tag{3.149}$$

It should be stressed that these are so-called "proper quasiparticle energies", representing the change in energy of the system *at fixed* N *and* R when the total magnetic flux is changed by a single Dirac quantum. The gap $E_{\mathrm{g}} \equiv \tilde{\varepsilon}^{\mathsf{QH}} + \tilde{\varepsilon}^{\mathsf{QP}} = (0.099 \pm 0.009)\, e^2/\epsilon\ell$ compares well with computations of Haldane and Rezayi (1985), who obtained $E_{\mathrm{g}} = (0.105 \pm 0.005)\, e^2/\epsilon\ell$, extrapolated from finite size exact diagonalization results.

A few words on how the sausage is made: Morf and Halperin worked with the Laughlin ground state and quasihole/quasielectron wavefunctions, evaluating the Coulomb energy for the system on the plane. Recall that the radius R of a Laughlin droplet at filling $\nu = q^{-1}$ is given by $R = \sqrt{2qN}\,\ell$. When a quasihole is created at the origin at fixed N, a small bubble is blown in the droplet, as each single particle angular momentum state is effectively shifted from $|\, m \,\rangle$ to $|\, m+1 \,\rangle$ (with some consequential changes in normalization[36]), hence the outer radius of the

[36] Note that since $\varphi_m(\boldsymbol{r}) = (2\pi\ell^2 m!)^{-1/2} (z/\sqrt{2}\ell)^m \exp(-|z|^2/4\ell^2)$ that $z\,\varphi_m(\boldsymbol{r}) = \sqrt{2(m+1)\ell}\,\varphi_{m+1}(\boldsymbol{r})$, and the j-dependence means that this multiplication is not precisely equivalent to adiabatically increasing the j quantum number.

droplet increases slightly. Since the largest single particle angular momentum is $m = q(N-1)$, in the state with a single quasihole at the origin the outer droplet radius shifts according to

$$R = \sqrt{2q(N-1)}\,\ell \quad \rightarrow \quad R = \sqrt{2[q(N-1)+1]}\,\ell \quad . \tag{3.150}$$

In order to keep the droplet radius fixed at the original value of $R = \sqrt{2q(N-1)}$, Morf and Halperin adjust the magnetic field, increasing it by a factor $1+[q(N-1)]^{-1}$, which has the effect of decreasing ℓ by just the right amount to prevent the outer radius of the droplet from shifting. Similarly for the quasielectron states, the field is reduced by a factor $1-[q(N-1)]^{-1}$.

MacDonald and Girvin[37] proposed the trial quasiparticle states

$$\begin{aligned} |\,\Psi_q^{\mathsf{QH}}\,\rangle &= \hat{U}\,|\,\Psi_q\,\rangle \\ |\,\Psi_q^{\mathsf{QE}}\,\rangle &= (1-\nu)^{-1/2}\,\hat{D}\,|\,\Psi_q\,\rangle \quad , \end{aligned} \tag{3.151}$$

where $\hat{U}\,|\,m\,\rangle = |\,m+1\,\rangle$ and $\hat{D}\,|\,m\,\rangle = (1-\delta_{m,0})\,|\,m-1\,\rangle$ raise and lower the single particle angular momentum quantum number of each electron, respectively[38]. Their quasiparticle energies were somewhat larger than those of Morf and Halperin:

$$\tilde{\varepsilon}^{\mathsf{QH}}_{\mathrm{GM}} = (0.0287 \pm 0.001)\,\frac{e^2}{\epsilon\ell} \qquad , \qquad \tilde{\varepsilon}^{\mathsf{QP}}_{\mathrm{GM}} = (0.085 \pm 0.002)\,\frac{e^2}{\epsilon\ell} \quad . \tag{3.152}$$

The energy gap E_{g} is related to the discontinuity in the chemical potential,

$$\Delta\mu = \left.\frac{\partial E}{\partial N}\right|_{\nu^+} - \left.\frac{\partial E}{\partial N}\right|_{\nu^-} = \frac{\tilde{\varepsilon}^{\mathsf{QE}} + \tilde{\varepsilon}^{\mathsf{QH}}}{|e^*/e|} = qE_{\mathrm{g}} \quad , \tag{3.153}$$

where $e^* = \pm e/q$. A pristine system should exhibit thermally activated resistivity according to $\rho_{xx}(T) \propto \exp(-E_{\mathrm{g}}/2k_{\mathrm{B}}T)$.

Adiabatic calculation of fractional quasiparticle charge and statistics

Consider a Hamiltonian $H(\boldsymbol{\lambda})$ where $\boldsymbol{\lambda} = \{\lambda_1, \ldots, \lambda_K\}$ are a set of parameters. Recall the definition of the geometric (Berry) connection

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\lambda}) = i\,\langle\,\Psi(\boldsymbol{\lambda})\,|\,\boldsymbol{\nabla}_{\boldsymbol{\lambda}}\,|\,\Psi(\boldsymbol{\lambda})\,\rangle \tag{3.154}$$

and the geometric phase

$$\gamma(\mathcal{C}) = \oint_{\mathcal{C}} d\boldsymbol{\lambda}\cdot\boldsymbol{\mathcal{A}}(\boldsymbol{\lambda}) \quad . \tag{3.155}$$

[37] A. H. MacDonald and S. M. Girvin, *Phys. Rev. B* **33**, 4414 (1986).

[38] The $(1-\nu)^{-1/2}$ factor in the quasielectron wavefunction is necessary to preserve normalization.

Recall also the Laughlin quasihole and quasielectron wavefunctions,

$$\begin{aligned}\Psi^{\rm QH}_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N;\boldsymbol{\xi}) &= \mathcal{M}(\boldsymbol{\xi})\prod_{l=1}^N(z_l-\xi)\prod_{j>k}^N(z_j-z_k)^q\prod_{i=1}^N\exp\bigl(-z_i\bar z_i/4\ell^2\bigr)\\ \Psi^{\rm QE}_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N;\boldsymbol{\eta}) &= \mathcal{N}(\boldsymbol{\eta})\prod_{i=1}^N\exp\bigl(-z_i\bar z_i/4\ell^2\bigr)\prod_{l=1}^N\left(2\ell^2\,\frac{\partial}{\partial z_l}-\bar\eta\right)\prod_{j>k}^N(z_j-z_k)^q\quad ,\end{aligned}\tag{3.156}$$

where $\mathcal{M}(\boldsymbol{\xi})$ and $\mathcal{N}(\boldsymbol{\eta})$ are normalization constants, which without loss of generality may be assumed to be real functions of their arguments. Treating the quasihole and quasielectron coordinates as adiabatic parameters, we may compute the geometric phase accrued as they each traverse a closed loop in two-dimensional space. Taking the differential of the quasihole wavefunction, we have

$$d\,\Psi^{\rm QH}(\boldsymbol{\xi}) = \Bigl[d\ln\mathcal{M}(\boldsymbol{\xi})+\sum_{i=1}^N d\ln(z_i-\xi)\Bigr]\,\Psi^{\rm QH}(\boldsymbol{\xi})\quad ,\tag{3.157}$$

where for notational simplicity we write $\Psi^{\rm QH}_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N;\boldsymbol{\xi})\equiv\Psi^{\rm QH}[\boldsymbol{\xi}]$. With $\boldsymbol{\lambda}=\boldsymbol{\lambda}(t)$ along the path $\mathcal{C}$, we compute the differential $d\gamma=\dot\gamma\,dt$, and find

$$d\gamma^{\rm QH} = i\,d\ln\mathcal{M}(\boldsymbol{\xi}) + i\,\langle\,\Psi^{\rm QH}(\boldsymbol{\xi})\,|\,\sum_{i=1}^N d\ln(z_i-\xi)\,|\,\Psi^{\rm QH}(\boldsymbol{\xi})\,\rangle\quad .\tag{3.158}$$

The number density in the quasihole state is

$$n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r}) = \langle\,\Psi^{\rm QH}(\boldsymbol{\xi})\,|\,\sum_{i=1}^N\delta(\boldsymbol{r}-\boldsymbol{r}_i)\,|\,\Psi^{\rm QH}(\boldsymbol{\xi})\,\rangle\quad .\tag{3.159}$$

Since the normalization $\mathcal{M}(\boldsymbol{\xi})$ is a real single-valued function of its argument, it cannot contribute to the integral for $\gamma(\mathcal{C})$, since it is the same at the initial and end points of any closed path. We now write $n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r}) = n+\delta n_{\boldsymbol{\xi}}(\boldsymbol{r})$, where $n=\nu/2\pi\ell^2$ is the density in the Laughlin ground state with $\nu=q^{-1}$ and $\delta n_{\boldsymbol{\xi}}(\boldsymbol{r})$ is concentrated about the location $\boldsymbol{\xi}$ of the quasihole defect. From the plasma analogy, we expect that in the thermodynamic limit that $\delta_n(\boldsymbol{\xi})$ should be a function of $|\boldsymbol{r}-\boldsymbol{\xi}|$ decaying on the scale of the Debye screening length $\ell/\sqrt{2}$. Therefore we have

$$\gamma^{\rm QH}(\mathcal{C}) = i\oint\limits_{|\xi|=R}d\xi\int d^2r\,\frac{n+\delta n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r})}{\xi-z}\tag{3.160}$$

Integrating $\boldsymbol{\xi}(t)$ over a circle of radius R, we have

$$\oint\limits_{|\xi|=R}\frac{d\xi}{\xi-z} = 2\pi i\,\Theta\bigl(R-|x|\bigr)\quad ,\tag{3.161}$$

where $\Theta(x)$ is a step function. Thus the background density term in $n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r})$ yields a contribution to the total Berry phase of

$$\gamma_0^{\rm QH}(\mathcal{C}) = i\int d^2r\; 2\pi i\, n\,\Theta(R-r) = -2\pi\,\langle N\rangle_{\mathcal{C}} = -2\pi\nu\,\Phi(\mathcal{C})/\phi_0 \quad , \tag{3.162}$$

where $\Phi(\mathcal{C}) = \pi R^2 B = \phi_0\, R^2/2\ell^2$ is the total magnetic flux enclosed by the loop $\mathcal{C}$. This is consistent with a Bohm-Aharonov phase of a charge $e^*_{\rm QH} = \nu e$ quasihole. The $\delta n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r})$ term integrates to zero if $\delta n^{\rm QH}_{\boldsymbol{\xi}}(\boldsymbol{r}) = \delta n\big(|\boldsymbol{r}-\boldsymbol{\xi}|\big)$ is a rotationally symmetric function in the difference $\boldsymbol{r}-\boldsymbol{\xi}$. This will be exponentially accurate if $\boldsymbol{\xi}$ lies in the bulk of the Laughlin droplet, far from the edge. Thus the full Berry phase is $\gamma^{\rm QH}(C) = -2\pi\nu\,\Phi(\mathcal{C})/\phi_0$.

To determine the statistics of the quasihole, we consider the state with two quasiholes:

$$\Psi^{\rm QH}(\boldsymbol{\xi},\boldsymbol{\xi}') = \mathcal{M}(\boldsymbol{\xi},\boldsymbol{\xi}')\prod_{l=1}^N (z_l-\xi)(z_l-\xi')\prod_{j>k}^N (z_j-z_k)^q\prod_{i=1}^N \exp\big(-z_i\bar z_i/4\ell^2\big) \quad . \tag{3.163}$$

We now carry out the same adiabatic calculation of the Berry phase by taking $\boldsymbol{\xi}$ around a circle of radius R, with $\boldsymbol{\xi}'$ held fixed. If $\boldsymbol{\xi}'$ lies outside the circle $|\boldsymbol{\xi}| = R$ by a distance d which is greater than just a few magnetic lengths, then the above analysis is unchanged, and the phase is $\gamma^{\rm QH}(\mathcal{C}) = -2\pi\nu\,\Phi(\mathcal{C})/\phi_0$. If, on the other hand, $\boldsymbol{\xi}'$ lies inside the loop, then there is a deficit in $\langle N\rangle_{\mathcal{C}}$ of $(-\nu)$, and the accrued phase is $\gamma'(\mathcal{C}) = \gamma^{\rm QH}(\mathcal{C}) + 2\pi\nu$, which says that when one quasihole winds around another, the wavefunction accumulates an extra *statistical phase* $\Delta\gamma(\mathcal{C}) = 2\pi\nu$. For exchange, one need only traverse half a circle, *i.e.* the relative angle of $\boldsymbol{\xi}$ and $\boldsymbol{\xi}'$ changes by π, leading to the *statistical angle* $\theta = \pi\nu$ for exchange of quasiholes. At $\nu = 1$, the statistical angle is $\theta_{\rm QH} = \pi$, corresponding to Fermi statistics, but for $\nu = q^{-1}$, the statistical angle is $\theta_{\rm QH} = \pi/q$, corresponding to *fractional statistics*.

Fractional statistics for particles in $d = 2$ was first discussed by Leinaas and Myrheim[39]. An interpretation of particles obeying fractional statistics in terms of charge-flux composites was first discussed by Wilczek[40], who called such particles *anyons*, because they could exhibit any type of exchange statistics. It was Halperin who first suggested that FQH quasiparticles obey fractional statistics, and argued that condensation of gases of these anyonic quasiparticles gave rise to a hierarchy of new FQH states at fillings $\nu = p/q$ with $p \neq 1$ and thus outside the principal Laughlin sequence of states. The adiabatic calculation of quasiparticle charge was first carried out by Arovas, Schrieffer, and Wilczek[41] (ASW).

The adiabatic calculation for quasielectrons is a tricky affair. The adiabatic argument of ASW analyzed

$$d\gamma^{\rm QE} = i\,d\ln\mathcal{N}(\boldsymbol{\eta}) + i\,\langle\,\Psi^{\rm QE}(\boldsymbol{\eta})\,|\,e^{-G(Z)}\left(\sum_{i=1}^N d\ln\big(2\ell^2\partial_i - \bar\eta\big)\right)e^{G(Z)}\,|\,\Psi^{\rm QE}(\boldsymbol{\eta})\,\rangle \quad , \tag{3.164}$$

[39] J. M. Leinaas and J. Myrheim, *Il Nuovo Cimento* **37**, 1, 1977.
[40] F. Wilczek, *Phys. Rev. Lett.* **48**, 1144 (1982).
[41] D. P. Arovas, J. R. Schrieffer, and F. Wilczek, *Phys. Rev. Lett.* **53**, 722 (1984).

where $G(Z) = \frac{1}{4\ell^2}\sum_{j=1}^N |z_j|^2$, appealing to the Girvin-Jach replacement $2\ell^2\partial_i \leftrightarrow \bar{z}_i$. This would appear to give

$$\gamma^{\mathsf{QE}}(\mathcal{C}) = i \oint\limits_{|\eta|=R} d\bar{\eta} \int d^2r \, \frac{n + \delta n^{\mathsf{QE}}_{\boldsymbol{\eta}}(\boldsymbol{r})}{\bar{\eta} - \bar{z}} \quad . \tag{3.165}$$

The $\bar{\eta}$ integral is taken clockwise around the path $\mathcal{C}$, hence this calculation yields the opposite value of the charge, *i.e.* $e^*_{\mathsf{QE}} = -\nu e$ for the quasielectron. Since the quasielectron thus represents an *increase* in electron density at $\boldsymbol{r} = \boldsymbol{\eta}$, there are two cancelling factors of (-1) in the calculation, and the statistical angle for quasielectrons is also $\theta_{\mathsf{QE}} = \pi/q$. Numerical calculations using Laughlin wavefunctions at $\nu = \frac{1}{3}$ for up to 200 particles were effected by Kjønsberg and Myrheim[42], who found good convergence for the adiabatic quasihole charge and statistics, but surprising poor convergence for the quasielectron values (especially the quasielectron statistical angle). These authors also found that boundary effects complicate the glib application of the Girvin-Jach replacement in the ASW calculations for the quasielectron.

The adiabatic method for determining effective quasiparticle charge and statistics is based on the adiabatic effective Lagrangian prescription of Moody, Shapere, and Wilczek[43], which goes as follows. Let $\Psi[\boldsymbol{\lambda}]$ be an adiabatic wavefunction and $\boldsymbol{\lambda}$ the adiabatic parameters. The adiabatic Lagrangian $L(\boldsymbol{\lambda}, \dot{\boldsymbol{\lambda}})$ is then given by

$$L(\boldsymbol{\lambda}, \dot{\boldsymbol{\lambda}}) = -\left\langle \Psi[\boldsymbol{\lambda}] \,\middle|\, \left\{ i\frac{d}{dt} + H(\boldsymbol{\lambda}) \right\} \,\middle|\, \Psi[\boldsymbol{\lambda}] \right\rangle = -\boldsymbol{\mathcal{A}}(\boldsymbol{\lambda}) \cdot \frac{d\boldsymbol{\lambda}}{dt} - E(\boldsymbol{\lambda}) \quad , \tag{3.166}$$

where $\boldsymbol{\mathcal{A}}(\boldsymbol{\lambda})$ is the Berry connection from §1.7.1. In the Born-Oppenheimer approach, where the nuclear coordinates $\{\boldsymbol{R}_1, \ldots, \boldsymbol{R}_{N_{\rm nuc}}\}$ are regarded as adiabatically varying so far as the electrons are concerned, the total effective Lagrangian is

$$L\big(\{\boldsymbol{R}_j\}, \{\dot{\boldsymbol{R}}_j\}\big) = \sum_{i=1}^{N_{\rm nuc}} \tfrac{1}{2} M \dot{\boldsymbol{R}}_i^2 - \sum_{i=1}^{N_{\rm nuc}} \boldsymbol{\mathcal{A}}_i\big(\{\boldsymbol{R}_j\}\big) \cdot \frac{d\boldsymbol{R}_i}{dt} - E_{\rm elec}\big(\{\boldsymbol{R}_j\}\big) - E_{\rm nuc}\big(\{\boldsymbol{R}_j\}\big) \quad . \tag{3.167}$$

3.3.7 Excitons

An exciton is a neutral entity formed from a quasielectron-quasihole pair bound by their mutually attractive Coulomb force. A callow description consists of charges $\pm e^* = \pm e/q$ located at positions $\boldsymbol{r}_{1,2}$. One then defines center-of-mass and relative coordinates $\boldsymbol{R} = \frac{1}{2}(\boldsymbol{r}_1 + \boldsymbol{r}_2)$ and $\boldsymbol{r} = \boldsymbol{r}_1 - \boldsymbol{r}_2$. Within the LLL, these objects have no inertial mass. If we assign each a mass m, the Lagrangian of the exciton system in the symmetric gauge becomes

$$L = \mathsf{m}\dot{\boldsymbol{R}}^2 + \tfrac{1}{4}\mathsf{m}\dot{\boldsymbol{r}}^2 - \frac{e^*B}{2c}\,\hat{\boldsymbol{z}} \cdot (\boldsymbol{R} \times \dot{\boldsymbol{r}} + \boldsymbol{r} \times \dot{\boldsymbol{R}}) - v(\boldsymbol{r}) \quad , \tag{3.168}$$

[42] H. Kjønsberg and J. Myrheim, *Int. Jour. Mod. Phys. A* **14**, 537, 1999.

[43] J. Moody, A. Shapere, and F. Wilczek, *Adiabatic Effective Lagrangians*, in *Geometric Phases in Physics* F. Wilczek and A. Shapere, eds. (World ScienRtific, 1989).

where $v(\boldsymbol{r}) = -e^{*2}/\epsilon r$ is the Coulomb interaction[44]. Thus

$$\begin{aligned} \boldsymbol{P} &= \frac{\partial L}{\partial \dot{\boldsymbol{R}}} = 2\mathsf{m}\dot{\boldsymbol{R}} + \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{r} \\ \boldsymbol{p} &= \frac{\partial L}{\partial \dot{\boldsymbol{r}}} = \tfrac{1}{2}\mathsf{m}\dot{\boldsymbol{r}} - \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{R} \quad . \end{aligned} \tag{3.169}$$

The equations of motion are

$$\begin{aligned} &2\mathsf{m}\ddot{\boldsymbol{R}} + \frac{e^*B}{c}\,\hat{\boldsymbol{z}}\times\dot{\boldsymbol{r}} = 0 \\ &\tfrac{1}{2}\mathsf{m}\ddot{\boldsymbol{r}} - \frac{e^*B}{c}\,\hat{\boldsymbol{z}}\times\dot{\boldsymbol{R}} = -\boldsymbol{\nabla} v(\boldsymbol{r}) \quad , \end{aligned} \tag{3.170}$$

and therefore

$$\frac{d}{dt}\Big(2\mathsf{m}\dot{\boldsymbol{R}} + \frac{e^*B}{c}\,\hat{\boldsymbol{z}}\times\boldsymbol{r}\Big) = 0 \quad , \tag{3.171}$$

Averaging over the fast cyclotron motion, we obtain

$$\langle\dot{\boldsymbol{R}}\rangle = -\frac{c}{e^*B}\,\hat{\boldsymbol{z}}\times\boldsymbol{\nabla}\tilde{v}(\boldsymbol{r}) \tag{3.172}$$

where $\tilde{v}(\boldsymbol{r})$ is an averaged potential as in §1.1.4 and $\langle\dot{\boldsymbol{r}}\rangle = 0$. Thus, the component charges of the exciton maintain their relative separation $\boldsymbol{r}$ and drift with speed $cv'(r)/e^*B$ in a direction perpendicular to $\boldsymbol{r}$.

The Hamiltonian of the exciton system is

$$H = \frac{1}{4\mathsf{m}}\Big(\boldsymbol{P} - \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{r}\Big)^2 + \frac{1}{\mathsf{m}}\Big(\boldsymbol{p} - \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{R}\Big)^2 + v(\boldsymbol{r}) \quad . \tag{3.173}$$

where $\boldsymbol{P}$ and $\boldsymbol{p}$ are the CM and relative coordinate canonical momenta, respectively. One may now define

$$\boldsymbol{G} = \sqrt{2}\,\Big(\boldsymbol{p} - \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{R}\Big) \qquad , \qquad \boldsymbol{\Lambda} = \frac{1}{\sqrt{2}}\,\Big(\frac{c}{e^*B}\,\hat{\boldsymbol{z}}\times\boldsymbol{P} + \tfrac{1}{2}\boldsymbol{r}\Big) \tag{3.174}$$

and

$$\boldsymbol{g} = \boldsymbol{P} + \frac{e^*B}{2c}\,\hat{\boldsymbol{z}}\times\boldsymbol{r} \qquad , \qquad \boldsymbol{\lambda} = -\frac{c}{e^*B}\,\hat{\boldsymbol{z}}\times\boldsymbol{p} + \tfrac{1}{2}\boldsymbol{R} \quad , \tag{3.175}$$

which satisfy

$$\big[G_\alpha\,,\,\Lambda_\beta\big] = -i\hbar\delta_{\alpha\beta} \qquad , \qquad \big[g_\alpha\,,\,\lambda_\beta\big] = -i\hbar\delta_{\alpha\beta} \quad , \tag{3.176}$$

which no other nonzero commutators. Then

$$H = \frac{\boldsymbol{G}^2}{2\mathsf{m}} + \tfrac{1}{2}\mathsf{m}\omega_{\rm c}^2\boldsymbol{\Lambda}^2 + v\Big(\sqrt{2}\,\boldsymbol{\Lambda} + \frac{c}{e^*B}\,\boldsymbol{g}\times\hat{\boldsymbol{z}}\Big) \quad , \tag{3.177}$$

[44] If $r \lesssim \ell$, the LLL projection blunts the $1/r$ divergence of the Coulomb interaction, as computed in §3.2.2.

where $\omega_{\rm c} = e^*B/{\rm m}c$. Thus $\big[H, \boldsymbol{g}\big] = 0$ and we may specify the momentum $\boldsymbol{g}$. Note that $\boldsymbol{g} = 2{\rm m}\dot{\boldsymbol{R}} + \frac{e^*B}{c}\,\hat{\boldsymbol{z}} \times \boldsymbol{r}$ was classically conserved. Recall that in charged systems at most one component of the total momentum could be fixed, *e.g.* in the Landau strip basis. Because the exciton is neutral, there is a conserved momentum. If we project the $\boldsymbol{G}$ and $\boldsymbol{\Lambda}$ degrees of freedom onto their lowest harmonic oscillator state, then we have $\big\langle \boldsymbol{\Lambda} \big\rangle = 0$ and $\big\langle \boldsymbol{\Lambda}^2 \big\rangle = \hbar c/e^*B$, and for large values of $\boldsymbol{g}$ we have that the energy is

$$\Delta_{\rm EX}(\boldsymbol{g}) = \tilde{\varepsilon}^{\rm QE} + \tilde{\varepsilon}^{\rm QH} + v\left(\frac{c}{e^*B}\,\boldsymbol{g} \times \hat{\boldsymbol{z}}\right) \quad . \tag{3.178}$$

Note $c/e^*B = q\ell^2/\hbar$ at filling $\nu = q^{-1}$.

3.3.8 Collective excitations

One might expect that there exist excited states of the FQH ground states corresponding to long-wavelength density oscillations, *i.e.* phonons. Whenever the ground state Ψ is of uniform density, it can often be argued on general grounds that such excitations are adequately represented by the *Ansatze*

$$|\,\Phi_{\boldsymbol{k}}\,\rangle = \frac{1}{\sqrt{N}} \sum_{i=1}^{N} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\,|\,\Psi\,\rangle = N^{-1/2}\rho_{\boldsymbol{k}}\,|\,\Psi\,\rangle \quad , \tag{3.179}$$

where

$$\rho_{\boldsymbol{k}} = \int d^2r\; e^{i\boldsymbol{k}\cdot\boldsymbol{r}} \overbrace{\sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i)}^{n(\boldsymbol{r})} = \sum_{i=1}^{N} e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i} \tag{3.180}$$

is the Fourier transform of the density. For $\boldsymbol{k} \to 0$, the state $|\,\Phi_{\boldsymbol{k}}\,\rangle$ will feature the same short-ranged correlations which favor the ground state $|\,\Psi\,\rangle$, yet $\langle\,\Psi\,|\,\Phi_{\boldsymbol{k}}\,\rangle = (2\pi)^2 N^{-1/2} n\,\delta(\boldsymbol{k})$ which vanishes for $\boldsymbol{k} \neq 0$. Thus, $|\,\Phi_{\boldsymbol{k}}\,\rangle$ serves as a trial state whose expected energy is a rigorous upper bound to the exact lowest excitation energy at wavevector $\boldsymbol{k}$.

The excitation energy of the state $|\,\Phi_{\boldsymbol{k}}\,\rangle$ is given by

$$\Delta(\boldsymbol{k}) = \frac{\langle\,\Phi_{\boldsymbol{k}}\,|\,H - E_0\,|\,\Phi_{\boldsymbol{k}}\,\rangle}{\langle\,\Phi_{\boldsymbol{k}}\,|\,\Phi_{\boldsymbol{k}}\,\rangle} = \frac{f(\boldsymbol{k})}{s(\boldsymbol{k})} \quad , \tag{3.181}$$

where

$$\begin{aligned} s(\boldsymbol{k}) &= \frac{1}{N}\,\langle\,\Psi\,|\,\rho^\dagger_{\boldsymbol{k}}\,\rho_{\boldsymbol{k}}\,|\,\Psi\,\rangle \\ f(\boldsymbol{k}) &= \frac{1}{2N}\,\langle\,\Psi\,|\,\big[\,\rho^\dagger_{\boldsymbol{k}}\,,\,[H, \rho_{\boldsymbol{k}}]\,\big]\,|\,\Psi\,\rangle \quad . \end{aligned} \tag{3.182}$$

The quantity $s(\boldsymbol{k})$ is the static structure factor, and $f(\boldsymbol{k})$ is known as the oscillator strength. When the Hamiltonian is of the form

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{p}_i^2}{2m} + \sum_{j<k} v(\boldsymbol{r}_j - \boldsymbol{r}_k) \quad , \tag{3.183}$$

the oscillator strength is given by

$$f(\boldsymbol{k}) = \frac{\hbar^2 \boldsymbol{k}^2}{2m} \quad , \tag{3.184}$$

independent of the potential v. This is known as the f-sum rule. It is also valid in a uniform magnetic field in which case $\boldsymbol{p}_i$ is replaced by $\boldsymbol{\pi}_i = \boldsymbol{p}_i + \frac{e}{c}\boldsymbol{A}(\boldsymbol{r}_i)$ for electrons. The static structure factor is

$$s(\boldsymbol{k}) = 1 + n \int d^2r \, \big[g(\boldsymbol{r}) - 1\big] \, e^{-i\boldsymbol{k}\cdot\boldsymbol{r}} + (2\pi)^2 n \, \delta(\boldsymbol{k}) \quad . \tag{3.185}$$

where the pair distribution function $g(\boldsymbol{r})$ is defined in Eqn. 3.112. Note that the last term above is not included in the definition of in Eqn. 3.113. This is because our definition in this section includes the diagonal $i = j$ term in what is the sum in Eqn. 3.112. This is a matter of convention and has no consequences for the following developments.

This approximation was originally employed by Feynman in deducing the phonon-roton spectrum of superfluid ^{4}He. It is really quite remarkable, for it allows one to represent a collective mode excitation spectrum solely in terms of static correlations in the ground state. In ^{4}He, $s(\boldsymbol{k})$ rises linearly[45] at small k and peaks at $k \equiv k_{\rm R} \propto n^{1/2}$, where n is the ground state number density. As a result, $\Delta(\boldsymbol{k})$ exhibits a local minimum at $k_{\rm R}$, called the *roton minimum*.

The dynamic structure factor (dsf) $S(\boldsymbol{k},\omega)$ is given by

$$S(\boldsymbol{k},\omega) = \frac{1}{N} \sum_j \big| \langle \, \Psi_j \, | \, \rho_{\boldsymbol{k}} \, | \, \Psi_0 \, \rangle \big|^2 \, \delta(\omega - \omega_j) \tag{3.186}$$

where $\hbar\omega_j \equiv E_j - E_0$. Here $|\,\Psi_0\,\rangle$ is the ground state and the sum is over the entire many-body spectrum of states $|\,\Psi_j\,\rangle$. We then have

$$\begin{aligned} s(\boldsymbol{k}) &= \int_0^\infty d\omega \, S(\boldsymbol{k},\omega) \\ f(\boldsymbol{k}) &= \int_0^\infty d\omega \, S(\boldsymbol{k},\omega) \, \hbar\omega \quad , \end{aligned} \tag{3.187}$$

and he have that $\Delta(\boldsymbol{k})$ is the first moment of the dsf. The expression for $\Delta(\boldsymbol{k})$ is therefore exact if $S(\boldsymbol{k},\omega)$ as a function of ω for each $\boldsymbol{k}$ has no variance, *i.e.* if $S(\boldsymbol{k},\omega) = S_{\rm SMA}(\boldsymbol{k},\omega)$, where

$$S_{\rm SMA}(\boldsymbol{k},\omega) = s(\boldsymbol{k}) \, \delta\big(\omega - \hbar^{-1}\Delta_{\rm SMA}(\boldsymbol{k})\big) \quad . \tag{3.188}$$

[45] See the Appendix §3.7.

Thus, $\Delta(\boldsymbol{k})$ is exact if a *single mode* saturates all the oscillator strength at wavevector $\boldsymbol{k}$. For this reason, this procedure is known as the *single mode approximation* or SMA. Precisely, *the SMA energy* $\Delta_{\rm SMA}(\boldsymbol{k})$ *is the exact first moment of the dynamic structure factor at wavevector* $\boldsymbol{k}$.

If one naïvely applies the SMA to the Laughlin ground state, a disappointing result is found. Invoking the result of Eqn. 3.123 for $\boldsymbol{k} \neq 0$, and with the f-sum rule fixing $f(\boldsymbol{k})$, we obtain

$$\Delta_{\rm SMA}(\boldsymbol{k}) = \hbar\omega_{\rm c}\left[1 - \tfrac{1}{4}(q-1)\,k^2\ell^2 + \mathcal{O}(k^4\ell^4)\right] \quad , \tag{3.189}$$

i.e. the excitation energy as $\boldsymbol{k} \to 0$ is independent of filling fraction and the interaction potential, and lies outside the lowest Landau level, at energy $\hbar\omega_{\rm c}$. The problem here is that the density operator $\rho_{\boldsymbol{k}}$ creates a mixture of inter-LL and intra-LL excitations, and for small k almost all of the oscillator strength is saturated by the inter-LL piece. This is known as *Kohn's theorem*[46], which says that the cyclotron resonance mode in systems without disorder saturates the oscillator strength up to terms of order k^2 as $k \to 0$. For any interaction $v_{ij} = v(\boldsymbol{r}_i - \boldsymbol{r}_j)$, the N-electron Hamiltonian

$$H = \sum_{i=1}^{N} \frac{\boldsymbol{\pi}_i^2}{2m^*} + \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j) \tag{3.190}$$

satisfies $[\,H, \Pi^\alpha\,] = i\hbar\omega_{\rm c}\,\epsilon_{\alpha\beta}\,\Pi^\beta$, with $\Pi^\alpha = \sum_i \pi_i^\alpha$. In particular, note that $[\,V, \Pi^\alpha\,] = 0$. Thus, $\Pi = \Pi^x + i\Pi^y$ is an eigenoperator of H, satisfying $[\,H, \Pi\,] = \hbar\omega_{\rm c}\,\Pi$. This means that if $|\,\Psi_0\,\rangle$ is the exact ground state of H, with $H\,|\,\Psi_0\,\rangle = E_0\,|\,\Psi_0\,\rangle$, then defining

$$|\,\Psi_1\,\rangle \equiv \frac{\Pi\,|\,\Psi_0\,\rangle}{\langle\,\Psi_0\,|\,\Pi^\dagger\,\Pi\,|\,\Psi_0\,\rangle^{1/2}} \quad , \tag{3.191}$$

we have $H\,|\,\Psi_1\,\rangle = (E_0 + \hbar\omega_{\rm c})\,|\,\Psi_1\,\rangle$, *i.e.* $|\,\Psi_1\,\rangle$ is an exact excited state with excitation energy $\hbar\omega_{\rm c}$ above the ground state, *independent of the interaction potential*. This is the Kohn mode.

To get at the intra-LL collective mode, Girvin, MacDonald, and Platzman[47] realized that what was needed is to project the density operator $\rho_{\boldsymbol{k}}$ onto the LLL. This is easily accomplished:

$$\bar\rho_{\boldsymbol{k}} = \Pi_0 \sum_{i=1}^{N} e^{ik\bar z_i/2}\, e^{i\bar k z_i/2}\, \Pi_0 = \sum_{i=1}^{N} e^{ik\ell b_i/\sqrt{2}}\, e^{i\bar k\ell b_i^\dagger/\sqrt{2}} \quad , \tag{3.192}$$

where we work in the symmetric gauge and invoke Eqn. 1.48, which expresses the complex coordinate z in terms of the cyclotron and guiding center ladder operators: $z = \sqrt{2}\,\ell(a + b^\dagger)$. Recall also the definition of the single particle magnetic translation operator,

$$t(\boldsymbol{d}) = e^{(db - \bar d b^\dagger)/\sqrt{2}\,\ell} \quad . \tag{3.193}$$

[46] W. Kohn, *Phys. Rev.* **123**, 142 (1961).
[47] S. M Girvin, A. H. MacDonald, and P. M. Platzman, *Phys. Rev. B* **33**, 2381 (1986).

Thus, we may write

$$\bar{\rho}_{\boldsymbol{k}} = e^{-\boldsymbol{k}^2\ell^2/4} \sum_{i=1}^{N} t_i(\ell^2 \hat{\boldsymbol{z}} \times \boldsymbol{k}) \quad , \tag{3.194}$$

where $t_i(\boldsymbol{d})$ is the MTO for the $i^{\rm th}$ particle. The MTOs satisfy the relations

$$t(\boldsymbol{a})\, t(\boldsymbol{b}) = e^{i\hat{\boldsymbol{z}}\cdot\boldsymbol{a}\times\boldsymbol{b}/2\ell^2}\, t(\boldsymbol{a}+\boldsymbol{b}) = e^{i\hat{\boldsymbol{z}}\cdot\boldsymbol{a}\times\boldsymbol{b}/\ell^2}\, t(\boldsymbol{b})\, t(\boldsymbol{a}) \quad . \tag{3.195}$$

Thus the projected density operators $\bar{\rho}_{\boldsymbol{k}}$ form a closed Lie algebra, *viz.*

$$\big[\bar{\rho}_{\boldsymbol{k}}\,,\,\bar{\rho}_{\boldsymbol{k}'}\big] = \big(e^{\bar{k}k'\ell^2/2} - e^{k\bar{k}'\ell^2/2}\big)\,\bar{\rho}_{\boldsymbol{k}+\boldsymbol{k}'} \quad , \tag{3.196}$$

which is known as the Girvin-MacDonald-Platzman (GMP) algebra[48]. Another useful result may be derived based on these relations:

$$\Pi_0\,\rho_{\boldsymbol{k}}\,\rho_{\boldsymbol{k}'}\,\Pi_0 = \bar{\rho}_{\boldsymbol{k}}\,\bar{\rho}_{\boldsymbol{k}'} + \big(1 - e^{\bar{k}k'\ell^2/2}\big)\bar{\rho}_{\boldsymbol{k}+\boldsymbol{k}'} \tag{3.197}$$

The LLL-projected Hamiltonian, after dropping the constant $\frac{1}{2}N\hbar\omega_{\rm c}$ cyclotron energy per particle, may be written

$$\begin{aligned}\bar{H} &= \Pi_0 \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j)\,\Pi_0 = \frac{1}{2}\int\frac{d^2k}{(2\pi)^2}\,\hat{v}(\boldsymbol{k})\,\Pi_0\,\rho^\dagger_{\boldsymbol{k}}\,\rho_{\boldsymbol{k}}\,\Pi_0 \\ &= \frac{1}{2}\int\frac{d^2k}{(2\pi)^2}\,\hat{v}(\boldsymbol{k})\Big(\bar{\rho}^\dagger_{\boldsymbol{k}}\,\bar{\rho}_{\boldsymbol{k}} - Ne^{-\boldsymbol{k}^2\ell^2/2}\Big) \quad .\end{aligned} \tag{3.198}$$

Applying a projected version of the SMA, then, GMP defined the state

$$|\,\Phi_{\boldsymbol{k}}\,\rangle = \frac{\bar{\rho}_{\boldsymbol{k}}\,|\,\Psi_0\,\rangle}{\langle\,\Psi_0\,|\,\bar{\rho}^\dagger_{\boldsymbol{k}}\,\bar{\rho}_{\boldsymbol{k}}\,|\,\Psi_0\,\rangle^{1/2}} \quad , \tag{3.199}$$

where $|\,\Psi_0\,\rangle$ is the exact ground state. Treating $|\,\Psi_{\boldsymbol{k}}\,\rangle$ as a trial state, we have

$$\Delta_{\rm SMA}(\boldsymbol{k}) = \langle\,\Phi_{\boldsymbol{k}}\,|\,\bar{H} - E_0\,|\,\Phi_{\boldsymbol{k}}\,\rangle = \frac{\bar{f}(\boldsymbol{k})}{\bar{s}(\boldsymbol{k})} \quad , \tag{3.200}$$

where

$$\begin{aligned}s(\boldsymbol{k}) &= \frac{1}{N}\,\langle\,\Psi_0\,|\,\bar{\rho}^\dagger_{\boldsymbol{k}}\,\bar{\rho}_{\boldsymbol{k}}\,|\,\Psi_0\,\rangle \\ f(\boldsymbol{k}) &= \frac{1}{2N}\,\langle\,\Psi_0\,|\,\big[\,\bar{\rho}^\dagger_{\boldsymbol{k}}\,,\,[H,\bar{\rho}_{\boldsymbol{k}}]\,\big]\,|\,\Psi_0\,\rangle \quad .\end{aligned} \tag{3.201}$$

[48] Note that $\rho^\dagger_{\boldsymbol{k}} = \rho_{-\boldsymbol{k}}$ and $\bar{\rho}^\dagger_{\boldsymbol{k}} = \bar{\rho}_{-\boldsymbol{k}}$, and also that $\rho_{\boldsymbol{0}} = \bar{\rho}_{\boldsymbol{0}} = 1$.

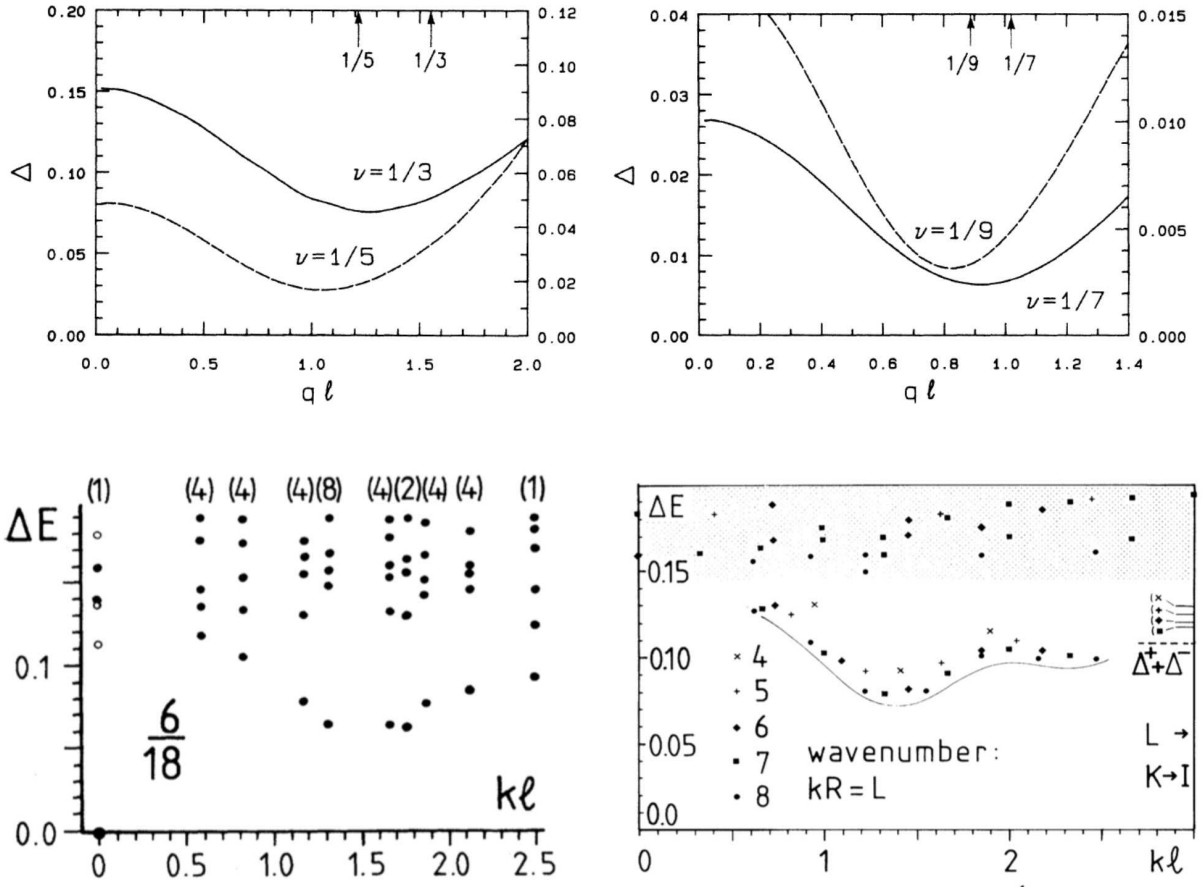


Figure 3.10: The theoretical and numerical dispersions for the magnetophonon-magnetoroton branch. Upper panels: Theoretical predictions of the collective mode dispersion at $\nu = \frac{1}{3}$, $\frac{1}{5}$, $\frac{1}{7}$, and $\frac{1}{9}$ (units of $e^2/\epsilon\ell$) from S. M. Girvin, A. M. MacDonald, and P. M. Platzman, *Phys. Rev. B* **33**, 2481 (1986). The arrows point to locations of the first reciprocal lattice vector of the corresponding Wigner crystal. Lower panels: Numerical computations of the excitation spectra for the Coulomb system at $\nu = \frac{1}{3}$ on the torus (left) and sphere (right), showing a clear $k \to 0$ gap and magnetoroton dip. From F. D. M. Haldane, *Phys. Rev. Lett.* **55**, 2095 (1985) and F. D. M. Haldane and E. H. Rezayi, *Phys. Rev. Lett.* **54**, 237 (1985).

are the projected structure factor and oscillator strength, which are given by

$$\begin{aligned} \bar{s}(\boldsymbol{k}) &= s(\boldsymbol{k}) - 1 + e^{-\boldsymbol{k}^2\ell^2/2} \\ \bar{f}(\boldsymbol{k}) &= \int \frac{d^2p}{(2\pi)^2}\, \hat{v}(\boldsymbol{p}) \left(1 - \cos(\ell^2 \hat{\boldsymbol{z}} \cdot \boldsymbol{k} \times \boldsymbol{p})\right) \left[\bar{s}(\boldsymbol{k}+\boldsymbol{p})\, e^{\ell^2 \boldsymbol{k}\cdot\boldsymbol{p}} - \bar{s}(\boldsymbol{p})\, e^{-\boldsymbol{k}^2\ell^2/2}\right] \quad . \end{aligned} \tag{3.202}$$

Appealing to Eqn. 3.123 and assuming this result, valid for the Laughlin states $|\,\Psi_q\,\rangle$, is also valid for the exact ground state $|\,\Psi_0\,\rangle$, we find

$$\bar{s}(\boldsymbol{k}) = \tfrac{1}{8}(q-1)\, k^4\ell^4 + \mathcal{O}(k^6\ell^6) \quad . \tag{3.203}$$

Note that the projected structure factor vanishes as k^4 in the long wavelength limit, in contrast to $s(\boldsymbol{k})$ itself, which vanishes as k^2. Now consider the lengthy expression for $\bar{f}(\boldsymbol{k})$ in the limit $\boldsymbol{k} \to 0$. The first factor in round brackets is proportional to k^2, The second factor in round brackets clearly vanishes when $\boldsymbol{k} = 0$, but the linear term in $\boldsymbol{k}$ must vanish after integrating over $\boldsymbol{p}$ whenever $v(\boldsymbol{p})$ and $\bar{s}(\boldsymbol{p})$ are isotropic. Thus the second factor in round brackets also vanishes as k^4 in the long wavelength limit. We conclude that the projected SMA results in the prediction of a *gap* in the collective mode spectrum at $\boldsymbol{k} = 0$. Note that when $q = 1$, the projected structure factor vanishes to all orders in k, because the SMA wavefunction itself vanishes (more on this below)!

Both the structure factor $s(\boldsymbol{k})$ and its projection $\bar{s}(\boldsymbol{k})$ exhibit a peak at the wavevector $k^* \approx \pi/a$ where $\pi a^2 n = 1$ with $n = \nu/2\pi\ell^2$. Thus $k^*\ell \approx \pi\sqrt{\nu/2}$. Although we don't have much intuition about the behavior of $\bar{f}(\boldsymbol{k})$, a natural guess is that the SMA energy $\Delta_{\text{SMA}}(\boldsymbol{k}) = \bar{f}(\boldsymbol{k})/\bar{s}(\boldsymbol{k})$ should exhibit a dip in the vicinity of k^*. Indeed this is what was found by GMP, whose results compared quite well with previous numerical studies of the excitation spectrum at $\nu = \frac{1}{3}$ by Haldane and by Haldane and Rezayi (1985). The $k \approx 0$ portion of this excitation branch is called the *magnetophonon*, and the $k \approx k^*$ portion the *magnetoroton*, using terminology borrowed from the study of superfluid ^{4}He. A comparison of theoretical predictions and numerical computations for this elementary excitation is shown in Fig. 3.10. Note that over a wide region centered at $k = k^*$, the magnetoroton appears as a true elementary excitation, *i.e.* as an isolated mode which contributes a δ-function to the dynamical structure factor $S(\boldsymbol{k}, \omega)$. While other states at $k \approx k^*$ are present at higher energies, they have much weaker oscillator strength, whence the accuracy of the SMA in this application.

GMP also examined the regime $k\ell \gg 1$, and found

$$\lim_{k\to\infty} \Delta_{\text{SMA}}(\boldsymbol{k}) = \frac{2\,\varepsilon_{\text{COH}}(\nu)}{1-\nu} \quad , \tag{3.204}$$

where $\varepsilon_{\text{COH}}(\nu) = \varepsilon(\nu) - \nu\varepsilon(1)$ is the cohesive energy per particle[49].

Note that while the SMA presumes that the underlying state $|\,\Psi_0\,\rangle$ is the exact ground state, GMP used the Laughlin states $|\,\Psi_q\,\rangle$ in its place. The agreement between the theoretical and numerically computed elementary excitation dispersion $\Delta_{\text{SMA}}(\boldsymbol{k})$ in the vicinity of its minimum is due to the accurate short-distance correlations encoded in the Laughlin state, and to the fact that the magnetoroton appears as a sharp collective mode in the excitation spectrum, and that this feature is not destroyed as one proceeds from the truncated pseudopotential Hamiltonian, for which the Laughlin state is exact, to the exact Coulomb ground state.

[49] The energy per particle is defined to be $\varepsilon(\nu) = E(\nu)/N$.

Charge susceptibility in the SMA

Suppose we add a dynamical perturbation to the Hamiltonian which couples to the density, described by

$$\bar{H}' = -\int d^2r\ \bar{\rho}(\boldsymbol{r})\, U(\boldsymbol{r},t) \quad . \tag{3.205}$$

Then from linear response theory,

$$\big\langle \delta\bar{\rho}(\boldsymbol{r},t)\big\rangle = \int\limits_{-\infty}^{\infty} dt' \int d^2r'\ \chi(\boldsymbol{r}-\boldsymbol{r}',t-t')\, U(\boldsymbol{r}',t') \quad , \tag{3.206}$$

where $\chi(\boldsymbol{r}-\boldsymbol{r}',t-t')$ is the *dynamical susceptibility*, whose Fourier transform is

$$\begin{aligned}\chi(\boldsymbol{k},\omega) &= \sum_j \big|\langle\, \Psi_0 \,|\, \bar{\rho}_{\boldsymbol{k}} \,|\, j\,\rangle\big|^2 \left\{\frac{1}{\omega+\omega_j+i\epsilon} - \frac{1}{\omega-\omega_j+i\epsilon}\right\} \\ &= \int\limits_{-\infty}^{\infty} d\nu\ S(\boldsymbol{k},\nu)\ \frac{2\nu}{\nu^2-(\omega+i\epsilon)^2} \quad .\end{aligned} \tag{3.207}$$

where $\epsilon = 0^+$ is a positive infinitesimal. Thus,

$$\delta\hat{\bar{\rho}}(\boldsymbol{k},\omega) = \chi(\boldsymbol{k},\omega)\,\hat{U}(\boldsymbol{k},\omega) \quad . \tag{3.208}$$

Replacing $S(\boldsymbol{k},\nu)$ in Eqn. 3.207 with the SMA result $\bar{s}(\boldsymbol{k})\,\delta\big(\hbar\nu - \Delta_{\text{SMA}}(\boldsymbol{k})\big)$, we obtain

$$\chi_{\text{SMA}}(\boldsymbol{k},\omega) = \frac{2\,\bar{s}(\boldsymbol{k})\,\Delta_{\text{SMA}}(\boldsymbol{k})}{\Delta^2_{\text{SMA}}(\boldsymbol{k}) - (\hbar\omega+i\epsilon)^2} \quad . \tag{3.209}$$

For static properties, at zero frequency, then, we have

$$\chi_{\text{SMA}}(\boldsymbol{k}) = \frac{2\bar{s}(\boldsymbol{k})}{\Delta_{\text{SMA}}(\boldsymbol{k})} = \frac{2\big[\bar{s}(\boldsymbol{k})\big]^2}{\bar{f}(\boldsymbol{k})} \quad . \tag{3.210}$$

Note that $\chi_{\text{SMA}}(\boldsymbol{k}) \sim k^4$ in the limit $k \to 0$. As an application, consider the response to an impurity potential $U(\boldsymbol{r})$. The resulting induced number density is $\langle\, \delta\hat{\bar{\rho}}(\boldsymbol{k})\,\rangle = \chi_{\text{SMA}}(\boldsymbol{k})\,\hat{U}(\boldsymbol{k})$. The total induced charge is obtained by taking the limit $\boldsymbol{k}\to 0$; this vanishes provided $\hat{U}(\boldsymbol{k})$ does not diverge as $k^{-\sigma}$ with $\sigma \geq 4$, where perturbation theory surely fails. The change in Coulomb energy of the electron system resulting from the perturbation is given by

$$\delta E_{\text{COUL}} = -\tfrac{1}{2}n\,\text{Re}\int\frac{d^2k}{(2\pi)^2}\,\hat{v}(-\boldsymbol{k})\,\langle\,\delta\hat{\bar{\rho}}(\boldsymbol{k})\,\rangle = -\tfrac{1}{2}n\,\text{Re}\int\frac{d^2k}{(2\pi)^2}\,\chi_{\text{SMA}}(\boldsymbol{k})\,\hat{v}(-\boldsymbol{k})\,\hat{U}(\boldsymbol{k}) \tag{3.211}$$

For a Coulomb impurity of charge $+Ze$, we have $\hat{U}(\boldsymbol{k}) = +Z\,\hat{v}(\boldsymbol{k})$ and the energy shift within the SMA is found to be

$$\delta E_{\text{COUL}} = -\tfrac{1}{2} Zn \int \frac{d^2k}{(2\pi)^2}\, \chi_{\text{SMA}}(\boldsymbol{k}) \left(\frac{2\pi e^2}{\epsilon k}\right)^2 \quad . \tag{3.212}$$

GMP report that evaluation of the above formula yields $\delta E_{\text{COUL}} = -1.15\, Ze^2/\epsilon\ell$ at $\nu = \frac{1}{3}$ (with $n = \nu/2\pi\ell^2$ as always), in substantial agreement with numerical calculations from exact diagonalization studies[50].

Orthogonality of SMA states of different wavevector

For the translationally invariant Bijl-Feynman wavefunction for ^{4}He, the density operator $\rho_{\boldsymbol{k}}$ changes the many-body momentum of the ground state. That is, if

$$T(\boldsymbol{d}) = \prod_{i=1}^{N} t_{0,i}(\boldsymbol{d}) = \prod_{i=1}^{N} e^{i\boldsymbol{p}_i \cdot \boldsymbol{d}/\hbar} \tag{3.213}$$

is the many-body translation operator, which implements $\boldsymbol{r}_i \to \boldsymbol{r}_i + \boldsymbol{d}$ for all $i \in \{1, \ldots, N\}$, then if the ground state is a state of definite momentum $\boldsymbol{P}_0$, then $T(\boldsymbol{d})\,|\,\Psi_0\,\rangle = e^{i\boldsymbol{P}_0\cdot\boldsymbol{d}/\hbar}\,|\,\Psi_0\,\rangle$ for any translation $\boldsymbol{d}$. It is then easy to prove that the state $|\,\Psi_{\boldsymbol{k}}\,\rangle = \rho_{\boldsymbol{k}}\,|\,\Psi_0\,\rangle$ is a state of momentum $\boldsymbol{P}_1 = \boldsymbol{P}_0 + \hbar\boldsymbol{k}$. This establishes that $\langle\,\Psi_{\boldsymbol{k}}\,|\,\Psi_{\boldsymbol{k}'}\,\rangle = 0$ for $\boldsymbol{k} \neq \boldsymbol{k}'$.

The situation is more complicated in the presence of a magnetic field, where the individual MTOs satisfy the algebra of Eqn. 3.215. If we place N electrons in a periodic (toroidal) geometry spanned by vectors $\boldsymbol{L}_{1,2}$ such that $\hat{\boldsymbol{z}} \cdot \boldsymbol{L}_1 \times \boldsymbol{L}_2 = 2\pi\ell^2 N_\phi$, then the many-body MTOs,

$$T(\boldsymbol{d}) = \prod_{i=1}^{N} t_i(\boldsymbol{d}) = \prod_{i=1}^{N} \exp\left(\frac{d\,b_i - \bar{d}\,b_i^\dagger}{\sqrt{2}\,\ell}\right) \tag{3.214}$$

satisfy the many-body version of Eqn. 3.215,

$$T(\boldsymbol{a})\,T(\boldsymbol{b}) = e^{iN\hat{\boldsymbol{z}}\cdot\boldsymbol{a}\times\boldsymbol{b}/2\ell^2}\,T(\boldsymbol{a}+\boldsymbol{b}) = e^{iN\hat{\boldsymbol{z}}\cdot\boldsymbol{a}\times\boldsymbol{b}/\ell^2}\,T(\boldsymbol{b})\,T(\boldsymbol{a}) \quad . \tag{3.215}$$

Then $T(\boldsymbol{L}_1)\,T(\boldsymbol{L}_2) = e^{2\pi i N N_\phi}\,T(\boldsymbol{L}_2)\,T(\boldsymbol{L}_1) = T(\boldsymbol{L}_2)\,T(\boldsymbol{L}_1)$ and therefore we can separately specify for each $a \in \{1, 2\}$ that $T(\boldsymbol{L}_a)\,|\,\Psi\,\rangle = e^{i\Theta_a}\,|\,\Psi\,\rangle$ for *all* states $|\,\Psi\,\rangle$ in our many-body Hilbert space. One can now check that

$$T(\boldsymbol{L}_a)\,t_i(\ell^2\hat{\boldsymbol{z}}\times\boldsymbol{k})\,T^\dagger(\boldsymbol{L}_a) = e^{i\boldsymbol{k}\cdot\boldsymbol{L}_a}\,t_i(\ell^2\hat{\boldsymbol{z}}\times\boldsymbol{k}) \tag{3.216}$$

[50] F. C. Zhang, V. Z. Vulovic, Y. Guo, and S. Das Sarma, *Phys. Rev. B* **32**, 6920 (1985) studied systems of up to $N = 4$ electrons in a toroidal geometry. The loss of translational invariance due to the impurity makes larger systems difficult to study numerically.

and therefore provided $e^{i\boldsymbol{k}\cdot\boldsymbol{L}_a} = 1$ all our SMA states lie in the same Hilbert space as specified by the eigenvalues of the unitaries $T(\boldsymbol{L}_{1,2})$. This establishes the proper quantization of the SMA wavevectors $\boldsymbol{k}$ in a periodic geometry, but it does not prove that SMA states of different wavevector are orthogonal.

We will delve more deeply into the many-body MTO algebra below. For now, let us establish the desired result by appealing to the thermodynamic limit for states on the plane. From Eqn. 3.197 we have, for any many-body state $|\,\Psi\,\rangle$ purely within the LLL,

$$\langle\,\Psi\,|\,\bar\rho^\dagger_{\boldsymbol{k}}\,\bar\rho_{\boldsymbol{k}'}\,|\,\Psi\,\rangle = \langle\,\Psi\,|\,\rho^\dagger_{\boldsymbol{k}}\,\rho_{\boldsymbol{k}'}\,|\,\Psi\,\rangle - \big(1 - e^{-\bar k k' \ell^2/2}\big)\langle\,\Psi\,|\,\rho_{\boldsymbol{k}'-\boldsymbol{k}}\,|\,\Psi\,\rangle \quad . \tag{3.217}$$

Now

$$\begin{aligned}\langle\,\Psi\,|\,\rho^\dagger_{\boldsymbol{k}}\,\rho_{\boldsymbol{k}'}\,|\,\Psi\,\rangle &= \int d^2r \int d^2r'\; n_2(\boldsymbol{r},\boldsymbol{r}')\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,e^{i\boldsymbol{k}'\cdot\boldsymbol{r}'} + \int d^2r\; n_1(\boldsymbol{r})\,e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}} \\ \langle\,\Psi\,|\,\rho_{\boldsymbol{k}'-\boldsymbol{k}}\,|\,\Psi\,\rangle &= \int d^2r\; n_1(\boldsymbol{r})\,e^{i(\boldsymbol{k}'-\boldsymbol{k})\cdot\boldsymbol{r}} \quad ,\end{aligned} \tag{3.218}$$

where

$$n_j(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_j) = \frac{N!}{(N-j)!}\int d^2x_{j+1}\cdots\int d^2x_N\;\big|\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_j,\boldsymbol{x}_{j+1},\ldots,\boldsymbol{x}_N)\big|^2 \tag{3.219}$$

are the diagonal elements of the j-particle density matrix. The j-particle distribution function is then defined as the ratio

$$g_j(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_j) \equiv \frac{n_j(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_j)}{n_1(\boldsymbol{r}_1)\cdots n_1(\boldsymbol{r}_j)} \quad . \tag{3.220}$$

In the thermodynamic limit, we may write $n_1(\boldsymbol{r}) = n$ and $n_2(\boldsymbol{r},\boldsymbol{r}') = n^2\,g(\boldsymbol{r}-\boldsymbol{r}')$, in which case

$$\langle\,\Psi\,|\,\bar\rho^\dagger_{\boldsymbol{k}}\,\bar\rho_{\boldsymbol{k}'}\,|\,\Psi\,\rangle = (2\pi)^2\left(n^2\,\hat h(\boldsymbol{k}) + n\,e^{-\boldsymbol{k}^2\ell^2/2}\right)\delta(\boldsymbol{k}-\boldsymbol{k}') + (2\pi)^4 n^2\,\delta(\boldsymbol{k})\,\delta(\boldsymbol{k}') \quad , \tag{3.221}$$

where $\hat h(\boldsymbol{k})$ is the Fourier transform of the pair correlation function. Note that in a system of finite area A we may replace $(2\pi)^2\,\delta(\boldsymbol{k}-\boldsymbol{k}') \to A\,\delta_{\boldsymbol{k}\boldsymbol{k}'}$. We thus have established the result

$$\langle\,\Phi_{\boldsymbol{k}}\,|\,\Phi_{\boldsymbol{k}'}\,\rangle = A\,\bar s(\boldsymbol{k})\,\delta_{\boldsymbol{k}\boldsymbol{k}'} + N^2\,\delta_{\boldsymbol{k},\boldsymbol{0}}\,\delta_{\boldsymbol{k}',\boldsymbol{0}} \tag{3.222}$$

where $|\,\Phi_{\boldsymbol{k}}\,\rangle = \bar\rho_{\boldsymbol{k}}\,|\,\Psi_0\,\rangle$, and where $\langle\,\Psi_0\,|\,\Psi_0\,\rangle = 1$ is assumed.

Final remarks on the SMA

In Eqn. 3.38 we saw how projecting a plane wave $e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}$ onto the LLL yields the result

$$\Pi_0\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,f(z)\,e^{-\boldsymbol{r}^2/4\ell^2} = e^{-\boldsymbol{k}^2\ell^2/2}\,e^{-i\bar k z/2}\,f(z - ik\ell^2)\,e^{-\boldsymbol{r}^2/4\ell^2} \quad . \tag{3.223}$$

The generalization to a many-body wavefunction $\Psi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) = P(z_1,\ldots,z_N)\,\prod_{i=1}^N e^{-\boldsymbol{r}_i^2/4\ell^2}$ is

$$\Pi_0\left(\prod_{i=1}^N e^{-\bar z_i z_i/4\ell^2}\right)\sum_{i=1}^N e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_i}\,P(z_1,\ldots,z_N) \tag{3.224}$$
$$= \left(\prod_{i=1}^N e^{-\bar z_i z_i/4\ell^2}\right)\underbrace{e^{-\boldsymbol{k}^2\ell^2/2}\sum_{j=1}^N P(z_1,\ldots,z_{j-1},z_j - ik\ell^2,z_{j+1},\ldots,z_N)\,e^{-i\bar k z_j/2}}_{\equiv\, P_{\boldsymbol{k}}(z_1,\ldots,z_N)}$$

Note that on the RHS, we have each z_j in $P_{\boldsymbol{k}}(Z)$ is in turn translated by $-ik\ell^2$. This messes with the zeroes of the wavefunction. In the Laughlin state, where $P(Z) = \big[V(Z)\big]^q$, the zeros of $P(Z)$ as a function of one of the holomorphic coordinates (say z_1) lie at the positions of all the other particles. In the SMA wavefunction, some of these zeros have shifted off the particle positions. However, since $\rho_{\boldsymbol{k}}$ is still symmetric under interchange of particle labels, we still have that $P_{\boldsymbol{k}}(Z_\sigma) = \mathsf{sgn}(\sigma)\,P_{\boldsymbol{k}}(Z)$. So at least $(N-1)$ of the $q(N-1)$ zeros like on the positions of the other particles. The remaining $(q-1)(N-1)$ zeros are shifted, and this results in an increase in energy, as is quite clear if we adopt a model truncated pseudopotential Hamiltonian.

What happens when $q = 1$ and $P(Z) = V(Z)$? Since $V(Z)$ is the only polynomial in our Hilbert space when $\nu = 1$, it must be that $\bar\rho_{\boldsymbol{k}}$ *annihilates* the filled LL whenever $\boldsymbol{k} \neq 0$. This is not so easy to show, however, and requires an appeal to the thermodynamic limit. Invoking the results of §3.1.2, we first obtain the second quantized form of the projected density operator in the angular momentum basis:

$$\bar\rho_{\boldsymbol{k}} = \sum_{m,m'} \langle\, m\,|\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,|\,m'\,\rangle\, c_m^\dagger\, c_{m'} \quad , \tag{3.225}$$

where the matrix element is

$$I_{m,m'}(\boldsymbol{k}) \equiv \langle\, m\,|\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,|\,m'\,\rangle = \int d^2r\;\varphi_m^*(\boldsymbol{r})\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}\,\varphi_{m'}(\boldsymbol{r}) \quad , \tag{3.226}$$

where $\varphi_m(\boldsymbol{r}) = (2\pi\ell^2 m!)^{-1/2}\,(z/\sqrt{2}\,\ell)^m\,\exp(-z\bar z/4\ell^2)$. We can use

$$I_{0,0}(\boldsymbol{k}) = \frac{1}{2\pi\ell^2}\int d^2r\;\; e^{-ik\bar z/2}\,e^{-i\bar k z/2}\,e^{-\bar z z/2\ell^2} = e^{-\bar k k\ell^2/2} \tag{3.227}$$

as a generating function, for which it is easily seen that

$$I_{m,m'}(\boldsymbol{k}) = \frac{1}{\sqrt{m!\,m'!}}\left(\frac{i\sqrt{2}}{\ell}\frac{\partial}{\partial k}\right)^m\left(\frac{i\sqrt{2}}{\ell}\frac{\partial}{\partial\bar k}\right)^{m'} e^{-\bar k k\ell^2/2} = \frac{i^{m-m'}}{\sqrt{m!\,m'!}}\left(\frac{\partial}{\partial w}\right)^m\big(w^{m'}e^{-\bar w w}\big) \tag{3.228}$$

with $w = k\ell/\sqrt{2}$, $\bar w = \bar k\ell/\sqrt{2}$, and $j = m' - m$. Thus,

$$I_{m,m+j}(\boldsymbol{k}) = (-i)^j\left(\frac{m!}{(m+j)!}\right)^{1/2}\left(\frac{k\ell}{\sqrt{2}}\right)^m L_m^{(j)}\big(\boldsymbol{k}^2\ell^2/2\big)\,e^{-\boldsymbol{k}^2\ell^2/2} \quad , \tag{3.229}$$

where $L_m^{(j)}(x)$ is a generalized Laguerre polynomial. Note $I_{m+j,m}(\boldsymbol{k}) = I^*_{m,m+j}(-\boldsymbol{k})$.

Now consider the action of the projected density operator $\bar\rho_{\boldsymbol{k}} = \sum_{m,m'} I_{m,m'}(\boldsymbol{k})\, c^\dagger_m\, c_{m'}$ on the filled Landau level $|\,\Psi_1\,\rangle = \prod_{m=0}^{N_\phi-1} c^\dagger_m\,|\,0\,\rangle$. In the thermodynamic limit, we have $N_\phi \to \infty$. Since every m state is occupied at $\nu = 1$, the off-diagonal terms in $\bar\rho_{\boldsymbol{k}}$ with $m \neq m'$ must annihilate $|\,\Psi_1\,\rangle$. This leaves only the diagonal elements, for which

$$I_{m,m}(\boldsymbol{k}) = L_m\big(\boldsymbol{k}^2\ell^2/2\big)\, e^{-\boldsymbol{k}^2\ell^2/2} \quad . \tag{3.230}$$

Thus,

$$\bar\rho_{\boldsymbol{k}}\,|\,\Psi_1\,\rangle = \sum_{m=0}^{\infty} I_{m,m}(\boldsymbol{k})\,|\,\Psi_1\,\rangle = e^{-\boldsymbol{k}^2\ell^2/2}\sum_{m=0}^{\infty} L_m(\boldsymbol{k}^2\ell^2/2)\,|\,\Psi_1\,\rangle \quad . \tag{3.231}$$

But the generating function of $L_n(x)$ yields[51]

$$\sum_{m=0}^{\infty} t^m\, L_m(x) = \frac{1}{1-t}\,\exp\!\left(-\frac{tx}{1-t}\right) \quad , \tag{3.232}$$

and thus taking $t \uparrow 1$ we encounter an essential singularity and provided $\boldsymbol{k} \neq 0$ the sum is zero! When $\boldsymbol{k} = 0$, from $L_m(0) = 1$ and cutting off the m sum at N_ϕ, we have $\bar\rho_{\boldsymbol{0}}\,|\,\Psi_1\,\rangle = N_\phi\,|\,\Psi_1\,\rangle$, which is also correct.

3.4 The Hierarchy

FQH plateaus have been observed at a number of odd-denominator rational fractions, including the principal Laughlin states at $\nu = \frac{1}{3}$ and $\nu = \frac{1}{5}$, their particle-hole conjugates at $\nu = \frac{2}{3}$ and $\nu = \frac{4}{5}$, and at a number of other fillings: $\nu = \frac{2}{5}$, $\frac{3}{5}$, $\frac{2}{7}$, $\frac{3}{7}$, $\frac{4}{7}$, $\frac{5}{7}$, $\frac{4}{9}$, $\frac{5}{11}$, $\frac{6}{13}$, *etc.* Other plateaus have been observed at $\nu = \frac{4}{3}$, $\frac{7}{5}$, $\frac{10}{7}$, *i.e.* in higher Landau levels. Fractions $\nu = p/q$ with $p \notin \{1, q-1\}$ are understood in terms of a hierarchical scheme originally due to Haldane and to Halperin, and a very successful set of wavefunctions due to Jain known as the composite fermion construction.

3.4.1 Particle-hole conjugation

Consider an arbitrary N-electron wavefunction within the LLL,

$$\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = P(z_1, \ldots, z_N)\,\exp\!\left(-\frac{1}{4\ell^2}\sum_{i=1}^{N}|z_i|^2\right) \quad , \tag{3.233}$$

[51] See Gradshteyn and Ryzhik 8.975.1.

how does one construct its particle-hole conjugate? Here we presume that $P(Z)$ is a multinomial function of its arguments. We start with the filled Landau level with N_ϕ total states, whose holomorphic component is the Vandermonde determinant $V(z_1, \ldots, z_{N_\phi})$, *viz.*

$$\Psi_1(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_{N_\phi}) = V(z_1, \ldots, z_{N_\phi}) \, \exp\left(-\frac{1}{4\ell^2} \sum_{i=1}^{N_\phi} |z_i|^2\right) \quad . \tag{3.234}$$

The unnormalized particle-hole conjugate wavefunction of the N-particle state in Eqn. 3.233 is constructed by taking its 'image' in the filled LL:

$$\Psi^{\rm C}(\boldsymbol{r}_{N+1}, \ldots, r_{N_\phi}) = \int d^2 r_1 \cdots \int d^2 r_N \; \Psi_1(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_{N_\phi}) \, \overline{\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N)} \quad , \tag{3.235}$$

where the bar denotes complex conjugation. This means that the holomorphic component of the M-particle wavefunction $\Psi^{\sf C}(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_M)$ is

$$P^{\sf C}(z_1, \ldots, z_M) = \int d^2 y_1 \cdots \int d^2 y_N \; V(y_1, \ldots, y_N, z_1, \ldots, z_M) \, \overline{P(y_1, \ldots, y_N)} \, \exp\left(-\frac{1}{2\ell^2} \sum_{i=1}^{N} |y_i|^2\right) \tag{3.236}$$

where $M + N = N_\phi$. Clearly $P^{\sf C}(Z)$ is totally antisymmetric in $Z = \{z_1, \ldots, z_M\}$. It is useful to introduce the notation $P[N] \equiv P(z_1, \ldots, z_N)$ to explicitly denote the number of holomorphic coordinates of P. We then have

$$\deg P^{\sf C}[M] = \tfrac{1}{2} N_\phi (N_\phi - 1) - \deg P[N] \quad , \tag{3.237}$$

and therefore if $\deg P[N] = N(N-1)/2\nu$, with $N = \nu N_\phi$, then

$$\deg P^{\sf C}[M] = \frac{N}{2\nu}\left(\frac{N}{\nu} - 1\right) - \frac{N}{2\nu}(N-1) = \frac{M(M-1)}{2(1-\nu)} \quad , \tag{3.238}$$

where $M = N_\phi - N = (\nu^{-1} - 1)N$. This of course confirms that the filling is

$$\nu^{\sf C} = \frac{M(M-1)}{2 \, \deg P^{\sf C}[M]} = 1 - \nu \quad . \tag{3.239}$$

3.4.2 Particle-hole symmetry

Consider the LLL-projected Coulomb Hamiltonian in the presence of a neutralizing background of density νn_0, where $n_0 = 1/2\pi\ell^2$:

$$H_\nu = \tfrac{1}{2} \int d^2 r \int d^2 r' \; v(\boldsymbol{r} - \boldsymbol{r}') \left\{ \psi^\dagger(\boldsymbol{r}) \, \psi^\dagger(\boldsymbol{r}') \, \psi(\boldsymbol{r}') \, \psi(\boldsymbol{r}) - 2\nu n_0 \, \psi^\dagger(\boldsymbol{r}) \, \psi(\boldsymbol{r}) + \nu^2 n_0^2 \right\} \quad . \tag{3.240}$$

Recall that the electron field operator is

$$\psi(\boldsymbol{r}) = \sum_\alpha \varphi_m(\boldsymbol{r})\, c_m \quad , \tag{3.241}$$

expressed in terms of the angular momentum basis (or any orthonormal basis complete in the LLL). Thus $\big\{\psi(\boldsymbol{r}), \psi^\dagger(\boldsymbol{r}')\big\} = n_0\, G(\boldsymbol{r}, \boldsymbol{r}')$ where $n_0 = 1/2\pi\ell^2$ and

$$G(\boldsymbol{r}, \boldsymbol{r}') = e^{z\bar{z}'/2\ell^2}\, e^{-z\bar{z}/4\ell^2}\, e^{-z'\bar{z}'/4\ell^2} = e^{i\hat{\boldsymbol{z}}\cdot\boldsymbol{r}\times\boldsymbol{r}'/2\ell^2}\, e^{-(\boldsymbol{r}-\boldsymbol{r}')^2/4\ell^2} = \big[G(\boldsymbol{r}', \boldsymbol{r})\big]^* \quad . \tag{3.242}$$

The anticommutation relations yield the following result,

$$\begin{aligned}\psi^\dagger(\boldsymbol{r})\,\psi^\dagger(\boldsymbol{r}')\,\psi(\boldsymbol{r}')\,\psi(\boldsymbol{r}) = \psi(\boldsymbol{r})\,\psi(\boldsymbol{r}')\,\psi^\dagger(\boldsymbol{r}')\,\psi^\dagger(\boldsymbol{r}) + n_0\,\psi^\dagger(\boldsymbol{r})\,\psi(\boldsymbol{r}) - n_0\,\psi(\boldsymbol{r}')\,\psi^\dagger(\boldsymbol{r}') \\ + n_0\,G(\boldsymbol{r}, \boldsymbol{r}')\,\psi(\boldsymbol{r}')\,\psi^\dagger(\boldsymbol{r}) - n_0\,G(\boldsymbol{r}', \boldsymbol{r})\,\psi^\dagger(\boldsymbol{r}')\,\psi(\boldsymbol{r}) \quad ,\end{aligned} \tag{3.243}$$

which establishes

$$H_\nu[\psi, \psi^\dagger] = H_{1-\nu}[\tilde{\psi}, \tilde{\psi}^\dagger] + \tfrac{1}{2}\, n_0 \int d^2r \int d^2r'\; v(\boldsymbol{r} - \boldsymbol{r}') \Big\{ G(\boldsymbol{r}, \boldsymbol{r}')\,\psi(\boldsymbol{r})\,\psi^\dagger(\boldsymbol{r}') - G(\boldsymbol{r}', \boldsymbol{r})\,\psi^\dagger(\boldsymbol{r})\,\psi(\boldsymbol{r}') \Big\} \tag{3.244}$$

where $\widetilde{\psi}(\boldsymbol{r}) = \psi^\dagger(\boldsymbol{r})$ and $\widetilde{\psi}^\dagger(\boldsymbol{r}) = \psi(\boldsymbol{r})$ is a particle-hole canonical transformation. Note however that $\big\{\widetilde{\psi}(\boldsymbol{r}), \widetilde{\psi}^\dagger(\boldsymbol{r}')\big\} = n_0\, G(\boldsymbol{r}', \boldsymbol{r})$. Of course this is is because particles and holes have opposite electric charge! We can now show that in any state $|\,\Psi_\nu\,\rangle$ in which $\langle\,\Psi_\nu\,|\,c_m^\dagger\, c_n\,|\,\Psi_\nu\,\rangle = \nu\,\delta_{mn}$, or equivalently $\langle\,\Psi_\nu\,|\,\psi^\dagger(\boldsymbol{r}')\,\psi(\boldsymbol{r})\,|\,\Psi_\nu\,\rangle = \nu\, G(\boldsymbol{r}, \boldsymbol{r}')$, that

$$\big\langle\,\Psi_\nu\,\big|\,H_\nu[\psi, \psi^\dagger]\,\big|\,\Psi_\nu\,\big\rangle + N_\phi\,\nu^2\sqrt{\frac{\pi}{8}}\,\frac{e^2}{\epsilon\ell} = \big\langle\,\widetilde{\Psi}_{1-\nu}\,\big|\,H_{1-\nu}[\tilde{\psi}, \tilde{\psi}^\dagger]\,\big|\,\widetilde{\Psi}_{1-\nu}\,\big\rangle + N_\phi\,(1-\nu)^2\sqrt{\frac{\pi}{8}}\,\frac{e^2}{\epsilon\ell} \quad , \tag{3.245}$$

where $|\,\widetilde{\Psi}_{1-\nu}\,\rangle \equiv |\,\Psi_\nu\,\rangle$. This establishes that if there is a cusp in the total energy at filling fraction ν, there will also be a cusp at filling fraction $1-\nu$.

3.4.3 Hierarchical construction of FQH wavefunctions

Haldane[52] was the first to suggest that Laughlin's quasiparticles could themselves condense into a higher order FQH state, which, in turn would itself have quasiparticle excitations which could condense, *ad infinitum*. Energetics would then determine how far along this chain one can proceed and still have stable condensates[53]. In Haldane's analysis, the quasiparticles carry bosonic statistics. Subsequently Halperin[54] devised a hierarchy based on fractional quasiparticle statistics, which we discussed in §3.3.6. Laughlin[55] proposed explicit hierarchical wavefunctions for the $\nu = \frac{2}{5}$ and $\frac{2}{7}$ states, and argued that the quasiparticles obeyed fermionic statistics.

[52] F. D. M. Haldane, *Phys. Rev. Lett.* **51**, 605 (1983).
[53] One competing phase would be a Wigner crystal of quasiparticles, for example.
[54] B. I. Halperin, *Phys. Rev. Lett.* **52**, 1982 (1984).
[55] R. B. Laughlin, *Surf. Sci.* **142**, 163 (1984).

Consistent with the Chern-Simons Ginzburg-Landau field theory of the FQHE, which we shall discuss in §CSGL below, I believe that a proper understanding of quasiparticle statistics in the FQHE entails that they are anyons, but the matter of their exchange statistics will not enter the following discussion.

To elicit hierarchical wavefunctions, we will follow the scheme of MacDonald, Aers, and Dharma-wardana[56], which relies heavily on the particle-hole conjugation formalism discussed in §3.4.1. In particular, recall the relation in Eqn. 3.236 between the N-particle holomorphic factor $P(z_1,\ldots,z_N)$ and its particle-hole conjugate $P^{\mathsf{C}}(z_1,\ldots,z_M)$. If at level $t-1$ of the hierarchy one has the holomorphic polynomial $P_{t-1}(Z)$ corresponding to a filling ν_{t-1}, we construct a new polynomial $P_t(Z)$ in one of two ways. The first way is to write

$$P_t(Z) = P^{\mathsf{C}}_{t-1}(Z)\,\big[V(Z)\big]^{2p_t} \tag{3.246}$$

with p_t a nonnegative integer. Note here that $P^{\mathsf{C}}_{t-1}(Z) = P^{\mathsf{C}}[N]$ is a function of the N holomorphic coordinates $\{z_1,\ldots,z_N\}$. Therefore, from Eqn. 3.238,

$$\begin{aligned}\deg P_t[N] = \frac{N(N-1)}{2\nu_t} &= \deg P^{\mathsf{C}}_{t-1}[N] + 2p_t \deg V[N] \\ &= \frac{N(N-1)}{2(1-\nu_{t-1})} + p_t\,N(N-1) \quad ,\end{aligned} \tag{3.247}$$

from which we obtain

$$\nu_t^{-1} = 2p_t + \frac{1}{1-\nu_{t-1}} = 2p_t + 1 + \frac{1}{\nu_{t-1}^{-1}-1} \quad . \tag{3.248}$$

Starting at level $i=0$ of the hierarchy with the polynomial $P_0(Z)\equiv 1$, which has $\deg P_0[N]=0$, corresponding to a filling $\nu_0 = 0$, the above formula then gives $\nu_1 = 1/(2p_1+1)$, which is one of the principal Laughlin states[57]. If we iterate the formula once more, with $p_2 = 0$, we obtain the particle-hole conjugate of the state at level $i=1$, with filling $\nu_2 = 2p_1/(2p_1+1)$.

The second iterative construction is given by[58]

$$P_t(Z) = V(Z)\left(\frac{P^{\mathsf{C}}_{t-1}(Z)}{V(Z)}\right)^{\dagger}\big[V(Z)\big]^{2p_t} \quad , \tag{3.249}$$

where $z_i^\dagger \equiv 2\ell^2\,\partial_t$. Note

$$\deg\left(\frac{P^{\mathsf{C}}_{t-1}[N]}{V[N]}\right) = \frac{N(N-1)}{2(1-\nu_{t-1})} - \frac{N(N-1)}{2} = \frac{N(N-1)}{2(\nu_{t-1}^{-1}-1)} \quad , \tag{3.250}$$

[56] A. H. MacDonald, G. C. Aers, and M. W. C. Dharma-wardana, *Phys. Rev. B* **31**, 5529 (1985).

[57] While the initial polynomial $P_0(Z)=1$ is not antisymmetric, this is not an issue. All the polynomials at level $i>0$ in the hierarchy are completely antisymmetric. We can think of $P_0(Z)$ as the limit as $\nu\to 0$ of a proper fermionic state.

[58] I think that another possible choice here would be to take $P_t(Z) = \big[P^{\mathsf{C}}_{t-1}(Z)\big]^\dagger\big[V(Z)\big]^{2p_t}$. This should yield the same fractions, but in with a different labeling.

and therefore

$$\deg P_t[Z] = \frac{N(N-1)}{2\nu_t} = \tfrac{1}{2}N(N-1) + p_t N(N-1) - \frac{N(N-1)}{2(\nu_{t-1}^{-1}-1)} \quad , \tag{3.251}$$

and thus

$$\nu_t^{-1} = 2p_t + 1 - \frac{1}{\nu_{t-1}^{-1} - 1} \quad . \tag{3.252}$$

We can express both steps of the hierarchy by the formula

$$\nu_t^{-1} = 2p_t + 1 + \frac{\alpha_{t-1}}{\nu_{t-1}^{-1} - 1} \quad , \tag{3.253}$$

where $\alpha_{t-1} \equiv +1$ if Eqn. 3.246 is used and $\alpha_{t-1} \equiv -1$ if Eqn. 3.249 is used. If we iterate the formulae t times, we obtain

$$\nu_t \equiv \big[p_t, \alpha_{t-1}\,p_{t-1}\,,\,\ldots\,,\,\alpha_0\,p_0\big] = \cfrac{1}{1 + 2p_t + \cfrac{\alpha_{t-1}}{2p_{t-1} + \cfrac{\alpha_{t-2}}{\ddots + \cfrac{\alpha_1}{2p_1 + \cfrac{\alpha_0}{2p_0}}}}} \tag{3.254}$$

It is convenient to write $\alpha p = \bar{p}$ for $\alpha = -1$. One then finds

$$\begin{aligned}
\tfrac{1}{3} &= [1] & \tfrac{2}{3} &= [0,1] \\
\tfrac{1}{5} &= [2] & \tfrac{4}{5} &= [0,2] \\
\tfrac{2}{5} &= [1,\bar{1}] & \tfrac{3}{5} &= [0,1,\bar{1}] \\
\tfrac{2}{7} &= [1,1] & \tfrac{5}{7} &= [0,1,1] \\
\tfrac{3}{7} &= [1,\bar{1},\bar{1}] & \tfrac{4}{7} &= [0,1,\bar{1},\bar{1}] \\
\tfrac{4}{9} &= [1,\bar{1},\bar{1},\bar{1}] & \tfrac{5}{9} &= [0,1,\bar{1},\bar{1},\bar{1}] \\
\tfrac{3}{11} &= [1,1,\bar{1}] & \tfrac{8}{11} &= [0,1,1,\bar{1}] \quad .
\end{aligned}$$

As is clear from the above, particle-hole conjugation means

$$\nu_t^{\mathsf{C}} = 1 - \nu_t = \big[0, p_t, \alpha_{t-1}\,p_{t-1}\,,\,\ldots\,,\,\alpha_0\,p_0\big] \quad . \tag{3.255}$$

Note that the same fraction may be represented by more than one sequence. For example,

$$\tfrac{4}{7} = [0,1,\bar{1},\bar{1}] = [1,1,\bar{1},\bar{1},\bar{1}] \quad , \tag{3.256}$$

which is a rather awkward aspect to this procedure. More seriously, one has to proceed rather deep into the hierarchy to arrive at several observed FQH fractions, such as $\nu = \frac{4}{9} = [1,\bar{1},\bar{1},\bar{1}]$.

3.4.4 Composite fermions

The composite fermion (CF) approach, pioneered by J. K. Jain[59], starts with the $\nu = r$ Slater determinant state of r filled LLs. In this state we have $N/N_\phi = r$ in the thermodynamic limit. We denote this state as Φ_r if $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$ as has been our convention, and as Φ_{-r} if $\boldsymbol{B} = +B\hat{\boldsymbol{z}}$. Now consider the flux attachment operation where $2p$ flux quanta are attached to each particle, yielding a *composite fermion*. This introduces $2p$ additional zeros whenever any two particles coincide, and is accomplished by multiplying the wavefunction $\Phi_{\pm r}$ for r filled LLs by the $(2p)^{\rm th}$ power of the Vandermonde determinant, yielding the N-particle wavefunction $\Psi_{\pm r,p}[N] = V^{2p}[N]\,\Phi_{\pm r}[N]$. The total flux per particle is then

$$\nu^{-1} = \frac{N_\phi}{N} = \pm\frac{1}{r} + 2p \qquad \Rightarrow \qquad \nu_{\pm r,p} = \frac{r}{2rp \pm 1} \quad . \tag{3.257}$$

Clearly this state will contain contributions from single-particle wavefunctions in the first n Landau levels, and if a *bona fide* LLL many-body wavefunction is desired, one should then *project* onto the $n = 0$ LL, *viz.*

$$\widetilde{\Psi}_{\pm r,p}[N] \equiv \Pi_0\,\Psi_{\pm r,p}[N] = \Pi_0\,V^{2p}[N]\,\Phi_{\pm r}[N] \quad . \tag{3.258}$$

Note that the projector Π_0 commutes with the total particle number N. Note also that the projector is applied *after* the state $\Phi_{\pm r}[N]$ is multiplied by $V^{2p}[N]$, since otherwise it is annihilated, *i.e.* $\Pi_0\,\Phi_{\pm r}[N] = 0$ for $n > 1$.

$\nu_{\pm r,p}$	$r=\bar{1}$	$r=1$	$r=\bar{2}$	$r=2$	$r=\bar{3}$	$r=3$	$r=\bar{4}$	$r=4$	$r=\bar{5}$	$r=5$
$p=0$	$\bar{1}$	1	$\bar{2}$	2	$\bar{3}$	3	$\bar{4}$	4	$\bar{5}$	5
$p=1$	1	1/3	2/3	2/5	3/5	3/7	4/7	4/9	5/9	5/11
$p=2$	1/3	1/5	2/7	2/9	3/11	3/13	4/15	4/17	5/19	5/21
$p=3$	1/5	1/7	2/11	2/13	3/17	3/19	4/23	4/25	5/29	5/31
$p=4$	1/7	1/9	2/15	2/17	3/23	3/25	4/31	4/33	5/39	5/41

Table 3.1: Filling fractions $\nu_{\pm r,p}$ for the first level of composite fermion states ($\bar{r} \equiv -r$). Observed fractions are printed in red. Not shown are the corresponding particle-hole conjugate states, for which $\nu^{\rm C}_{\pm r,p} = 1 - \nu_{\pm r,p}$.

One might worry about the effect of the projector, but Jain found that it has a rather weak effect,

[59] J. K. Jain, *Phys. Rev. Lett.* **63**, 199 (1989); J. K. Jain, *Phys. Rev. B* **41**, 7653 (1990). See also J. K. Jain, *Composite Fermions* (Cambridge, 2007).

and *most* of the state $\Psi_{\pm r,p}[N]$ is confined to the LLL, with

$$\begin{aligned}\nu = \tfrac{2}{5}\ (r=2,p=1):\ \ &\frac{\langle\,\Psi_{r,p}\,|\,\Pi_0\,|\,\Psi_{r,p}\,\rangle}{\langle\,\Psi_{r,p}\,|\,\Psi_{r,p}\,\rangle} \approx 0.05 \\ \nu = \tfrac{4}{9}\ (r=4,p=1):\ \ &\frac{\langle\,\Psi_{r,p}\,|\,\Pi_0\,|\,\Psi_{r,p}\,\rangle}{\langle\,\Psi_{r,p}\,|\,\Psi_{r,p}\,\rangle} \approx 0.01 \quad .\end{aligned} \tag{3.259}$$

With $\widetilde{\Psi}_{\pm r,p}[N]$ fully in the LLL, we can construct its particle-hole conjugate $\Psi^{\mathsf{C}}_{\pm r,p}[N]$, with filling $\nu^{\mathsf{c}}_{\pm r,p} = 1 - \nu_{\pm r,p}$. A table of values for small n and p is given in Tab. 3.1. States in this sequence may be thought of as describing an integer quantum Hall state of composite fermions. A touted success of the CF theory is that most of the observed states appear at early in the sequence.

Note that the $r \to \infty$ limit of this sequence yields a filling

$$\lim_{r\to\infty} \nu_{\pm r,p} = \frac{1}{2p} \quad , \tag{3.260}$$

which is an *even denominator* fermionic state with $r^{-1} = 0$ flux per composite fermion. If the wavefunction is $\Psi[N] = V^{2p}[N]\,\Phi[N]$, then the component $\Phi[N]$ is a state corresponding to zero flux. One possibility is a Slater determinant of plane waves, *i.e.*

$$\Phi(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) = A^{-N/2}\,\mathsf{det}\big(e^{i\boldsymbol{k}_i\cdot\boldsymbol{r}_j}\big) \quad , \tag{3.261}$$

If the wavevectors $\{\boldsymbol{k}_1,\ldots,\boldsymbol{k}_N\}$ are arranged in a Fermi circle, then there must be *gapless particle-hole excitations* corresponding to removing a state just below $k_{\rm F}$ and replacing it with one just above $k_{\rm F}$. We will discuss the theory of the half-filled Landau level in the next chapter.

The CF states described above represent only the first level of the CF hierarchy. At the second level, one can start, for instance, with the $\nu = \frac{4}{3}$ state, which is $\nu = \frac{1}{3}$ state in the $n = 1$ LL sitting atop a filled $n = 0$ LL[60]. Thus $\nu = N/N_\phi = \frac{4}{3}$ and $N_\phi = \frac{3}{4}N$. The flux attachment operation results in $\nu^{-1} \to \nu^{-2} + 2$ hence $\nu' = \nu/(2\nu+1)$ and with $\nu = \frac{4}{3}$ we obtain $\nu' = \frac{4}{11}$, which is an observed fraction that is in fact not present among the states of the first level of the CF hierarchy. The states $\frac{5}{13}$, $\frac{7}{11}$, $\frac{4}{13}$, $\frac{6}{17}$, and $\frac{5}{17}$ are also observed (see Fig. 3.11 and are present

[60]We assume 100% spin polarization, so we work only with spinless fermion wavefunctions here.

only at the second level of the hierarchy. Note that under flux addition we have[61]

$$\begin{aligned}
\frac{4}{3} &\longrightarrow \frac{4/3}{(8p/3)\pm 1} \underset{p=1}{=} \begin{cases} 4/11 \\ 4/5 \end{cases} \\
\frac{4}{3} &\longrightarrow \frac{4/3}{(8p/3)\pm 1} \underset{p=2}{=} \begin{cases} 4/19 \\ 4/13 \end{cases} \\
\frac{6}{5} &\longrightarrow \frac{6/5}{(12p/5)\pm 1} \underset{p=1}{=} \begin{cases} 6/17 \\ 6/7 \end{cases} \\
\frac{5}{3} &\longrightarrow \frac{5/3}{(10p/3)\pm 1} \underset{p=1}{=} \begin{cases} 5/13 \\ 5/7 \end{cases} \\
\frac{7}{3} &\longrightarrow \frac{7/3}{(14p/3)\pm 1} \underset{p=1}{=} \begin{cases} 7/17 \\ 7/11 \end{cases}
\end{aligned} \tag{3.262}$$

States in this sequence may be thought of as describing a fractional quantum Hall state of composite fermions. The complete set of operations generating the CF states is thus:

(i) Landau level addition: $\nu \to \nu + 1$

(ii) flux attachment: $\nu^{-1} \to \nu^{-1} + 2$

(iii) LLL projection (preserves N): $\Psi \to \Pi_0 \Psi$

(iv) particle-hole conjugation (LLL states only): $\nu \to 1 - \nu$.

3.5 Chern-Simons Ginzburg-Landau Theory

The understanding that the Laughlin states and their hierarchical descendants were especially stable gapped phases of matter with peculiar topological properties (*e.g.* FQHE, quasiparticle excitations obeying fractional exchange statistics) naturally led researchers to think of these states as some sort of novel condensates. If so, what is the corresponding order parameter and continuum field theory? In fact, today it is known that topological phases of matter such as the FQHE phases do not possess an order parameter of the usual sort and do not break any global symmetries. Nevertheless, a compelling field-theoretic description of the FQHE, due to Zhang, Hansson, and Kivelson[62], based in part on earlier work by Girvin and MacDonald[63] has proven extremely useful and influential. I especially recommend the pellucid review by Zhang[64].

[61] Red typeface indicates observed FQH fractions; blue typeface indicates fractions already present at the first level of the CF hierarchy.

[62] S.-C. Zhang, H. Hansson, and S. Kivelson, *Phys. Rev. Lett.* **62**, 82 (1989).

[63] S. M. Girvin and A. H. MacDonald, *Phys. Rev. Lett.* **58**, 1252 (1987).

[64] S.-C. Zhang, *Int. Jour. Mod. Phys. B* **6**, 25 (1992).

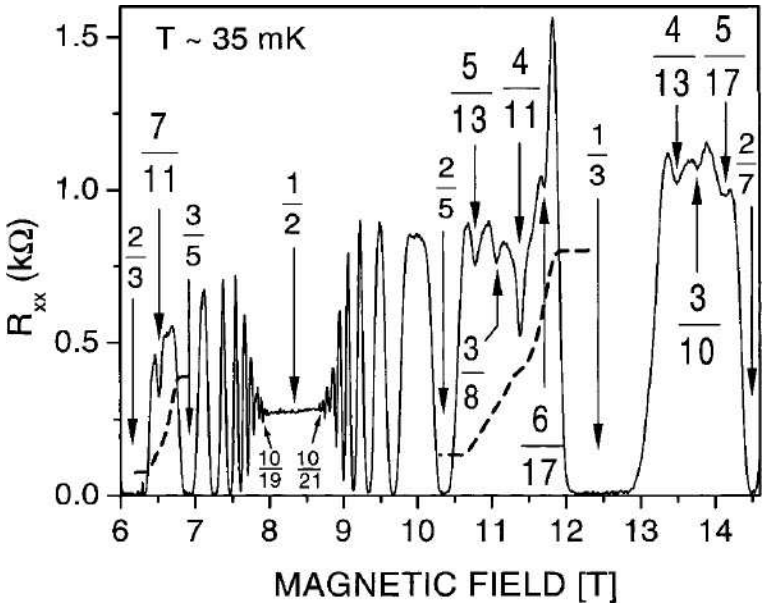


Figure 3.11: Experimental observation of second level composite fermion states: $\nu = \frac{4}{11}$, $\frac{5}{13}$, $\frac{7}{11}$, $\frac{4}{13}$, $\frac{6}{17}$, and $\frac{5}{17}$. From W. Pan *et al.*, *Phys. Rev. Lett.* **90**, 016801 (2003).

3.5.1 Superfluids, vortices, and duality

As a warm-up, we will first consider the case of a $2+1$-dimensional superfluid. The Chern-Simons Ginzburg-Landau (CSGL) theory establishes a connection between the field theory of the superfluid and that of the FQHE. Of course, superfluidity is a phenomenon of bosonic systems, whereas the constituent particles of the FQHE are electrons, which are fermions. In superconductors, the electrons first pair before condensing, as do the helium atoms in liquid ^{3}He[65]. But this is not what happens in the Laughlin states, for example. Rather, in order to describe the FQHE in terms of a bosonic field theory of an order parameter field, we will have to manufacture fermions from bosons – a trick known as *statistical transmutation*. As we shall see, at a mean field level, the implementation of the statistical transmutation, which is effected using a fictitious gauge field, can *cancel* with the physical magnetic field, leaving behind a purely bosonic theory in zero field, but with telltale fluctuation terms. The vortices of the bosonic superfluid then correspond to quasiparticles of the FQHE! But all good things to those who wait; first let's examine the case of superfluidity in $(2+1)$-dimensions[66].

[65]The fact that electrons are charged, whereas ^{3}He atoms are neutral, entails some essential corrections to the simplistic description of a superconductor as a Bose condensate.

[66]V. N. Popov, *Sov. Phys. JETP* **37**, 341 (1973); M. P. A. Fisher and D.-H. Lee, *Phys. Rev. B* **39**, 2756 (1989); G. E. Volovik, *JETP Lett.*, **62**, 65 (1995); E. Šimánek, *Inhomogeneous Superconductors* (Oxford, 1994); D. P. Arovas and J. A. Freire, *Phys. Rev. B* **55**, 1068 (1997).

Start with the Euclidean action

$$S_{\rm E} = \int d\tau \int d^2x \left\{ \hbar\,\bar\psi\,\partial_\tau\,\psi + \frac{\hbar^2}{2m}\left|\boldsymbol{\nabla}\psi\right|^2 \right\} + \tfrac{1}{2}\int d\tau \int d^2x \int d^2x'\; v(\boldsymbol{x}-\boldsymbol{x}')\,\delta n(\boldsymbol{x})\,\delta n(\boldsymbol{x}') \quad , \tag{3.263}$$

where

$$\delta n(\boldsymbol{x}) \equiv n(\boldsymbol{x}) - n_0 \quad . \tag{3.264}$$

We write $\psi = \sqrt{n}\,e^{i\phi}$, in which case

$$\begin{aligned} \partial_\tau\psi &= \left(\frac{\partial_\tau n}{2\sqrt{n}} + i\sqrt{n}\,\partial_\tau\phi\right) e^{i\phi}. \\ \left|\boldsymbol{\nabla}\psi\right|^2 &= \frac{\left(\boldsymbol{\nabla}n\right)^2}{4n} + n\left(\boldsymbol{\nabla}\phi\right)^2 \quad . \end{aligned} \tag{3.265}$$

Setting aside the interaction term for the moment, the Lagrangian density, other than the v term, is then given by

$$\begin{aligned} \mathcal{L}_0 &= i\hbar\,n\,\partial_\tau\phi + \frac{\hbar^2 n}{2m}\left(\boldsymbol{\nabla}\phi\right)^2 + \frac{\hbar^2}{8mn}\left(\boldsymbol{\nabla}n\right)^2 \\ &\underset{\rm HST}{\longrightarrow} i\hbar\,n\,\partial_\tau\phi + i\hbar\,\boldsymbol{Q}\cdot\boldsymbol{\nabla}\phi + \frac{m\,\boldsymbol{Q}^2}{2n} + \frac{\hbar^2}{8mn}\left(\boldsymbol{\nabla}n\right)^2 \quad , \end{aligned} \tag{3.266}$$

where $\boldsymbol{Q}$ is a Hubbard-Stratonovich field. We now separate $\phi = \phi_{\rm sw} + \phi_{\rm v}$ into the smooth spin-wave and singular vortex contributions. Integrating over $\phi_{\rm sw}$ yields the constraint

$$\partial_\tau n + \boldsymbol{\nabla}\cdot\boldsymbol{Q} = 0 \quad , \tag{3.267}$$

which is solved by writing

$$\begin{aligned} n - n_0 &= n_0\left(\frac{\partial\mathcal{W}_x}{\partial y} - \frac{\partial\mathcal{W}_y}{\partial x}\right) \equiv -n_0\,\mathcal{B} \\ Q_x &= n_0\left(\frac{\partial\mathcal{W}_y}{\partial\tau} - \frac{\partial\mathcal{W}_\tau}{\partial y}\right) \equiv -un_0\,\mathcal{E}_y \\ Q_y &= n_0\left(\frac{\partial\mathcal{W}_\tau}{\partial x} - \frac{\partial\mathcal{W}_x}{\partial\tau}\right) \equiv +un_0\,\mathcal{E}_x \quad , \end{aligned} \tag{3.268}$$

where we have defined the dimensionless analog electromagnetic fields $(\boldsymbol{\mathcal{E}}, \mathcal{B})$ and where u is an as-yet undetermined constant with dimensions of speed. We also define the background gauge feld $w_\mu(\boldsymbol{x},\tau) = -x\,\hat{\boldsymbol{y}}$. Regarding the vortex part, we have

$$i\hbar\,n\,\partial_\tau\phi_{\rm v} + i\hbar\,\boldsymbol{Q}\cdot\boldsymbol{\nabla}\phi_{\rm v} = -2\pi i\hbar n_0\,J^\mu\left(\mathcal{W}_\mu + w_\mu\right) + \partial\left(\,\cdot\,\right) \quad , \tag{3.269}$$

where

$$J^\mu(\boldsymbol{x},\tau) = \frac{1}{2\pi}\,\epsilon^{\mu\nu\lambda}\,\partial_\nu\partial_\lambda\phi_{\mathsf{v}} = \sum_i q_i \left\{ \begin{array}{c} 1 \\ \dot{\boldsymbol{X}}_i \end{array} \right\} \delta\big(\boldsymbol{x} - \boldsymbol{X}_i(\tau)\big) \quad , \tag{3.270}$$

where $q_i \in \mathbb{Z}$ is the integer vorticity of vortex i, and $x^\mu = (\tau, x, y)$ for $\mu = 0, 1, 2$. There is no difference between raised and lowered indices here. Thus, dropping total derivatives, we have

$$\mathcal{L}_0 = \frac{\hbar^2}{8mn}\,(\boldsymbol{\nabla} n)^2 + \frac{m\boldsymbol{Q}^2}{2n} - 2\pi i\hbar n_0\,J^\mu\left(\mathcal{W}_\mu + w_\mu\right) + \partial\,(\,\cdot\,) \tag{3.271}$$

We now expand to quadratic order in δn and $\boldsymbol{Q}$. Including the v term in the potential, we have

$$\begin{aligned} S_{\mathsf{E}} &= \int\frac{d\omega}{2\pi}\int\frac{d^2k}{(2\pi)^2}\left\{\frac{m}{2n_0}\left|\hat{\boldsymbol{Q}}(\boldsymbol{k},\omega)\right|^2 + \left(\tfrac{1}{2}\hat{v}(\boldsymbol{k}) + \frac{\hbar^2\boldsymbol{k}^2}{8mn_0}\right)\left|\delta\hat{n}(\boldsymbol{k},\omega)\right|^2\right\} - ihn_0\int d\tau\int d^2x\;J^\mu\left(\mathcal{W}_\mu + w_\mu\right) \\ &= \tfrac{1}{2}n_0mu^2\int\frac{d\omega}{2\pi}\int\frac{d^2k}{(2\pi)^2}\left(\left|\hat{\boldsymbol{\mathcal{E}}}(\boldsymbol{k},\omega)\right|^2 + \frac{\omega_{\mathsf{p}}^2(\boldsymbol{k})}{u^2\boldsymbol{k}^2}\left|\hat{\mathcal{B}}(\boldsymbol{k})\right|^2\right) - ihn_0\int d\tau\int d^2x\;J^\mu\left(\mathcal{W}_\mu + w_\mu\right) \end{aligned} \tag{3.272}$$

where $h = 2\pi\hbar$ and the phonon dispersion is

$$\omega_{\mathsf{p}}^2(\boldsymbol{k}) = \frac{n_0}{m}\,\boldsymbol{k}^2\,\hat{v}(\boldsymbol{k}) + \frac{\hbar^2|\boldsymbol{k}|^4}{4m^2} \quad . \tag{3.273}$$

Thus if $\hat{v}(\boldsymbol{0})$ is finite, we may define $u = \sqrt{n_0\hat{v}(\boldsymbol{0})/m}$, which is the phonon velocity at long wavelengths. The Euclidean Lagrangian density in the long wavelength limit is then

$$\mathcal{L}_{\mathsf{E}} = \tfrac{1}{2}n_0mu^2\big(\boldsymbol{\mathcal{E}}^2 - \mathcal{B}^2\big) - 2\pi i\hbar n_0 J^\mu(\mathcal{W}_\mu + w_\mu) \quad , \tag{3.274}$$

which is to say $(2+1)$-dimensional electrodynamics in the presence of a background magnetic field $b = \epsilon_{ij}\partial_i w_j = -1$. Note that the Fourier transform of the vortex 3-current is

$$\hat{J}^\mu(\boldsymbol{k},\omega) = \int d\tau\,\sum_i q_i \left\{ \begin{array}{c} 1 \\ \dot{\boldsymbol{X}}_i \end{array} \right\} e^{i\omega\tau}\,e^{-i\boldsymbol{k}\cdot\boldsymbol{X}_i(\tau)} \quad . \tag{3.275}$$

Note also that

$$-\,2\pi i\hbar n_0\int\limits_{-\infty}^{\infty}\!dt\;J^\mu\,w_\mu = 2\pi i\hbar n_0\sum_i q_i\int d\tau\;X_i(\tau)\,\dot{Y}_i(\tau) \tag{3.276}$$

gives the geometric phase for vortices winding about the condensate. If $\hat{v}(\boldsymbol{0})$ diverges, as is the case for the Coulomb potential $v(\boldsymbol{r}) = e^2/\epsilon r$, for which $\hat{v}(\boldsymbol{k}) = 2\pi e^2/\epsilon|\boldsymbol{k}|$, then there is no long-wavelength low-frequency effective Lorentz invariance with speed u. For $v(\boldsymbol{r}) = e^2/\epsilon r$, one has $\hat{v}(\boldsymbol{k}) = 2\pi e^2/\epsilon|\boldsymbol{k}|$ and the phonon disperses as $|\boldsymbol{k}|^{1/2}$ as $\boldsymbol{k} \to 0$.

Charged bosons

If the bosons have a charge $\tilde{e}$, this may be accommodated by the substitutions

$$\partial_\tau \phi_{\mathsf{v}} \to \partial_\tau \phi_{\mathsf{v}} + \tilde{e} A^0 \qquad , \qquad \boldsymbol{\nabla} \phi_{\mathsf{v}} \to \boldsymbol{\nabla} \phi_{\mathsf{v}} + \tfrac{\tilde{e}}{c} \boldsymbol{A} \quad , \tag{3.277}$$

where A^μ is the electromagnetic vector potential and c is the speed of light. Note that $c \gg u$. One may now define

$$K^\mu \equiv J^\mu + \frac{\tilde{e}}{2\pi} \epsilon^{\mu\nu\lambda} \partial_\nu A_\lambda = \begin{cases} J^0 - \tilde{e} B / 2\pi c \\ J^x - \tilde{e} E_y / 2\pi \\ J^y + \tilde{e} E_x / 2\pi \end{cases} , \tag{3.278}$$

where $A^\mu \equiv (A^0, c^{-1} \boldsymbol{A})$, and replace J^μ by K^μ, also including the Maxwell term $(\boldsymbol{E}^2 - B^2)\, d / 8\pi$ in the (2+1)-dimensional Euclidean Lagrangian density, where d is the thickness of the system.

Integrating out the gauge field

Next we integrate out the gauge field $\mathcal{W}_\mu$. We define the fields $\mathcal{W}_\parallel$ and $\mathcal{W}_\perp$ by

$$\begin{aligned} \mathcal{W}_x(\boldsymbol{k}, \omega) &= i \hat{k}_x \, \mathcal{W}_\parallel(\boldsymbol{k}, \omega) - i \hat{k}_y \, \mathcal{W}_\perp(\boldsymbol{k}, \omega) \\ \mathcal{W}_y(\boldsymbol{k}, \omega) &= i \hat{k}_y \, \mathcal{W}_\parallel(\boldsymbol{k}, \omega) + i \hat{k}_x \, \mathcal{W}_\perp(\boldsymbol{k}, \omega) \quad , \end{aligned} \tag{3.279}$$

where $\hat{\boldsymbol{k}} = \boldsymbol{k} / |\boldsymbol{k}|$. We then have

$$\delta \hat{n}(\boldsymbol{k}, \omega) = n_0 \big(i k_y \mathcal{W}_x - i k_x \mathcal{W}_y \big) = n_0 \, |\boldsymbol{k}| \, \mathcal{W}_\perp(\boldsymbol{k}, \omega) \tag{3.280}$$

and

$$\begin{aligned} \hat{Q}_x(\boldsymbol{k}, \omega) &= n_0 \big(- i \omega \mathcal{W}_y - i k_y \mathcal{W}_0 \big) \\ &= n_0 \big(- i k_y \mathcal{W}_0(\boldsymbol{k}, \omega) + \omega \hat{k}_y \mathcal{W}_\parallel(\boldsymbol{k}, \omega) + \omega \hat{k}_x \mathcal{W}_\perp(\boldsymbol{k}, \omega) \big) \end{aligned} \tag{3.281}$$

and

$$\begin{aligned} \hat{Q}_y(\boldsymbol{k}, \omega) &= n_0 \big(i \omega \mathcal{W}_x + i k_x \mathcal{W}_0 \big) \\ &= n_0 \big(i k_x \mathcal{W}_0(\boldsymbol{k}, \omega) - \omega \hat{k}_x \mathcal{W}_\parallel(\boldsymbol{k}, \omega) + \omega \hat{k}_y \mathcal{W}_\perp(\boldsymbol{k}, \omega) \big) \quad . \end{aligned} \tag{3.282}$$

Thus,

$$\begin{aligned} \big| \delta \hat{n}(\boldsymbol{k}, \omega) \big|^2 &= n_0^2 \, \boldsymbol{k}^2 \, \big| \mathcal{W}_\perp(\boldsymbol{k}, \omega) \big|^2 \\ \big| \hat{\boldsymbol{Q}}(\boldsymbol{k}, \omega) \big|^2 &= n_0^2 \, \omega^2 \, \big| \mathcal{W}_\perp(\boldsymbol{k}, \omega) \big|^2 + n_0^2 \, \Big| \, \omega \, \mathcal{W}_\parallel(\boldsymbol{k}, \omega) - i \, |\boldsymbol{k}| \, W_0(\boldsymbol{k}, \omega) \, \Big|^2 \quad . \end{aligned} \tag{3.283}$$

Next, we write

$$\begin{aligned}\int d\tau \int d^2x\; J^\mu\, \mathcal{W}_\mu &= \sum_i q_i \int d\tau \int \frac{d\omega}{2\pi} \int \frac{d^2k}{(2\pi)^2}\; e^{-i\omega\tau}\, e^{i\boldsymbol{k}\cdot\boldsymbol{X}_i(\tau)} \Big\{ i\hat{\boldsymbol{k}} \times \dot{\boldsymbol{X}}_i \cdot \hat{\boldsymbol{z}}\, \mathcal{W}_\perp(\boldsymbol{k},\omega) \\ &\qquad\qquad + \mathcal{W}_0(\boldsymbol{k},\omega) + \big(i\hat{k}_x \dot{X}_i + i\hat{k}_y \dot{Y}_i\big)\, \mathcal{W}_\parallel(\boldsymbol{k},\omega) \Big\} \\ &= \sum_i q_i \int d\tau \int \frac{d\omega}{2\pi} \int \frac{d^2k}{(2\pi)^2}\; e^{-i\omega\tau}\, e^{i\boldsymbol{k}\cdot\boldsymbol{X}_i(\tau)} \Big\{ i\hat{\boldsymbol{k}} \times \dot{\boldsymbol{X}}_i \cdot \hat{\boldsymbol{z}}\, \mathcal{W}_\perp(\boldsymbol{k},\omega) \\ &\qquad\qquad + \frac{i}{|\boldsymbol{k}|}\Big(\omega\, \mathcal{W}_\parallel(\boldsymbol{k},\omega) - i\,|\boldsymbol{k}|\, \mathcal{W}_0(\boldsymbol{k},\omega)\Big) \Big\} \quad . \end{aligned} \tag{3.284}$$

In obtaining the last line above we have used

$$\big(i\hat{k}_x \dot{X}_i + i\hat{k}_y \dot{Y}_i\big)\, e^{i\boldsymbol{k}\cdot\boldsymbol{X}_i(\tau)} = \frac{1}{|\boldsymbol{k}|}\, \frac{\partial}{\partial\tau}\, e^{i\boldsymbol{k}\cdot\boldsymbol{X}_i(\tau)} \quad , \tag{3.285}$$

and then integrated by parts. We now integrate out the fields $\mathcal{W}_\perp(\boldsymbol{k},\omega)$ and the combination $\omega\,\mathcal{W}_\parallel(\boldsymbol{k},\omega) - i|\boldsymbol{k}|\,\mathcal{W}_0(\boldsymbol{k},\omega)$. Integrating out the latter yields a frequency-independent kernel $|\boldsymbol{k}|^{-2}$ and an instantaneous logarithmic Coulomb interaction among the vortices. Thus,

$$\begin{aligned} S_{\mathsf{E}} &= 2\pi i\hbar n_0 \sum_i q_i \int d\tau\, X_i(\tau)\, \dot{Y}_i(\tau) - \frac{\pi\hbar^2 n_0}{m} \sum_{i,j} q_i\, q_j \int d\tau\; \ln\big|\boldsymbol{X}_i(\tau) - \boldsymbol{X}_j(\tau)\big| \\ &\quad + \frac{\pi\hbar^2 n_0}{m} \sum_{i,j} q_i\, q_j \int d\tau \int d\tau' \int \frac{d\omega}{2\pi} \int \frac{d^2k}{(2\pi)^2}\; \frac{\hat{\boldsymbol{k}} \wedge \dot{\boldsymbol{X}}_i(\tau)\; \hat{\boldsymbol{k}} \wedge \dot{\boldsymbol{X}}_j(\tau')}{\omega^2 + \omega_{\mathsf{p}}^2(\boldsymbol{k})}\; e^{-i\omega(\tau-\tau')}\, e^{i\boldsymbol{k}\cdot\left(\boldsymbol{X}_i(\tau) - \boldsymbol{X}_j(\tau')\right)} \quad , \end{aligned} \tag{3.286}$$

where $\boldsymbol{a} \wedge \boldsymbol{b} \equiv \boldsymbol{a} \times \boldsymbol{b} \cdot \hat{\boldsymbol{z}}$.

Thus, we end up with a theory of logarithmically interacting vortices, whose dynamics are equivalent to those of electrons in the lowest Landau level due to the $X_i\,\dot{Y}_i$ term in the Lagrangian, plus a retarded interaction described by the last term. This is $(2+1)$-dimensional electrodynamics, where vortices play the role of charges, and phonons play the role of photons.

3.5.2 Statistical transmutation

One usually does not think of quantum statistics as a continuous parameter, like a coupling constant. We are accustomed to the notion that many-particle wavefunctions are either symmetric or antisymmetric, *i.e.* $\Psi(\cdots j \cdots i \cdots) = e^{i\theta}\,\Psi(\cdots i \cdots j \cdots)$ with $\theta = 2n\pi$ for bosons and $\theta = (2n+1)\pi$ for fermions. Other values of θ, such as $\theta = \frac{1}{2}\pi$, seem to make no sense because iterating the relation twice gives $\Psi(\cdots i \cdots j \cdots) = e^{2i\theta}\,\Psi(\cdots i \cdots j \cdots)$, and thus $e^{2i\theta} \neq 1$ contradicts single-valuedness of Ψ. One concludes that Bose and Fermi statistics exhaust all possible values of θ.

What happens, though, if we relax the single-valuedness constraint and allow the wavefunction $\Psi\big(\{\boldsymbol{r}_j\}\big)$ to be a *multivalued* function of its arguments? One example of a multivalued function is the complex function $f(z) = z^\alpha$, which changes by a factor $e^{2\pi i\alpha}$ when z moves counterclockwise around a circle enclosing the origin. Paths which wind around the origin n times accumulate a phase factor of $e^{2\pi in\alpha}$. If α is not an integer, then $f(z)$ returns to its original value multiplied by a phase. Although it may seem strange to consider multivalued wavefunctions, there is nothing that prevents us from doing so. The Schrödinger equation is a differential equation and thus only requires that Ψ be *locally* well-defined. In addition, physical quantities, such as probability densities, always depend on $|\Psi|^2$ and are appropriately single-valued, as the multivaluedness we consider will always be in the *phase* of the wavefunction.

In the example $f(z) = z^\alpha$, z takes its values in the complex plane. In the case of many-particle quantum mechanics, the argument $\boldsymbol{R} \equiv \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\}$ of $\Psi(\boldsymbol{R})$ lives in a more complicated space, called *configuration space*. It is the space of all N-tuples $\boldsymbol{R}$ together with the equivalence relation $\boldsymbol{R} \cong \sigma\boldsymbol{R}$, where $\sigma \in \mathcal{S}_N$ is any element of the permutation group, so that $\sigma\boldsymbol{R} = \{\boldsymbol{r}_{\sigma(1)}, \ldots, \boldsymbol{r}_{\sigma(N)}\}$. The equivalence of $\boldsymbol{R}$ and $\sigma\boldsymbol{R}$ means that the particles are indistinguishable. For technical purposes, it is necessary to impose the restriction that no two particles ever lie at the same position – this is necessary for the multivaluedness to be meaningful. This is analogous to the situation in our simple example of $f(z) = z^\alpha$ above, in which paths that intersect the origin cannot be assigned a definite winding number. Physically, the restriction that no two particles lie atop one another can be accomplished by imposing an infinitely repulsive hard core potential of vanishingly small range; this has no effect on any physical properties.

We now ask what sorts of multivalued functions can be defined on this configuration space. Recall that in the case of the simple example $f(z) = z^\alpha$ that paths could be classified by an integer winding number n; paths which have the same winding number are equivalent to one another in the sense that they can be smoothly deformed into each other without crossing the origin. Associated to each path of winding number n was a phase $e^{2\pi in\alpha}$. If we append one path of winding number n' onto a path of winding number n, the resultant path has winding number $n + n'$. Thus, we can think of the space of paths as a mathematical group, and in this simple case, group addition of two paths of winding numbers n and n' produces a third path of winding number $n + n'$. Mathematically, this result is succinctly stated as

$$\pi_1\big(\mathbb{R}^2 \setminus \{\boldsymbol{0}\}\big) \cong \mathbb{Z} \quad , \tag{3.287}$$

which means that the group of paths (under the operation of path addition) on the punctured plane (the plane minus the origin) is isomorphic to the group of integers (under the operation of addition). Mathematicians refer to the group of paths $\pi_1(\mathcal{M})$ as the *fundamental group*, or 'first homotopy group' *first homotopy group* of the manifold $\mathcal{M}$. The fundamental group of the punctured plane is isomorphic to the integers.

The configuration space for N identical particles living on a base manifold $\mathcal{M}$ is

$$\mathcal{X}_N(\mathcal{M}) = \big(\mathcal{M}^N - \mathcal{D}\big)/\mathcal{S}_N \quad , \tag{3.288}$$

where $\mathcal{D} = \big\{(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) \mid \boldsymbol{r}_i = \boldsymbol{r}_j \text{ for some } i \neq j\big\}$. Note as with the punctured plane, we exclude certain subspaces from our manifold, in this case corresponding to coincidences of the positions of at least two particles. This space is more complicated than the punctured plane. The difference is that rather than classifying paths by how they wind around the origin, we classify paths by how the particles wind around other particles. If the base space $\mathcal{M}$ is d-dimensional, then $\dim \mathcal{X}_N(\mathcal{M}) = dN$. Consider a closed path in this configuration space from a point $\boldsymbol{R}$ to an equivalent point $\sigma\boldsymbol{R}$. If $d \geq 3$, it is easy to see that any two paths from $\boldsymbol{R}$ to $\sigma\boldsymbol{R}$ are deformable into one another. Just as loops in $\mathbb{R}^3$ cannot be classified by a winding number (they can be shrunk to a point without ever crossing the origin), any two configuration space paths $\boldsymbol{R}$ to $\sigma\boldsymbol{R}$ are *homotopically equivalent*, *i.e.* they can be deformed into one another. The paths are then classified by σ alone. The mathematicians would say that

$$\pi_1\big(\mathcal{X}_N(\mathcal{M})\big) \cong \mathcal{S}_N \quad (\dim(\mathcal{M}) > 2) \quad . \tag{3.289}$$

The phases associated with the paths form a unitary one-dimensional representation of the fundamental group $\pi_1\big(\mathcal{X}_N(\mathcal{M})\big)$, and so for $d = \dim(\mathcal{M}) > 2$, we are left with unitary one-dimensional representations of $\mathcal{S}_N$, of which there are only two: the symmetric (Bose) representation, $e^{i\theta_\sigma} = +1$, and the antisymmetric (Fermi) representation, $e^{i\theta_\sigma} = \operatorname{sgn}\sigma$.

In two space dimensions, the notion of relative winding of particles becomes well-defined. As a consequence, the space of loops in configuration space becomes more complicated. Indeed, a path in which a particle winds completely around another particle can no longer be deformed to a point without crossing that particle. The fundamental group of configuration space is no longer $\mathcal{S}_N$, but rather is an infinite nonabelian group, known as the N-string *braid group*[67] on $\mathcal{M}$, *i.e.* $\mathcal{B}_N(\mathcal{M})$. As its name suggests, the algebra of this group is associated with the weaving of 'braids', which are world-lines for our particles. The phases associated with the paths in configuration space now form a unitary one-dimensional representation of the braid group: to each pairwise exchange of particles one associates a factor $e^{i\theta}$. If we let $z_j = x_j + iy_j$ be the complex coordinate for particle j, the wavefunction takes the form

$$\Psi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N) = \prod_{i<j} (z_i - z_j)^{\theta/\pi}\, \Phi(\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N), \tag{3.290}$$

where $\Phi(\boldsymbol{R})$ is a totally symmetric function. Note that $\theta = \pi$ leads to a function which satisfies Fermi statistics[68]. Intermediate values of $\theta \in (0, \pi)$ correspond to fractional statistics, first clearly discussed by Leinaas and Myrheim in 1977[69]. The above configuration space analysis is due to Laidlaw and DeWitt[70], who were mostly concerned with $d = 3$, and to Wu[71], who considered the case $d = 2$ in detail.

[67] The braid groups were first introduced by E. Artin in 1928. For a review, see R. Fox and L. Neuwirth, *Math. Scand.* **10**, 119 (1962).

[68] It should be noted that conventional wavefunctions satisfying Fermi statistics *are* multivalued in configuration space, since their sign changes depending on the parity of the permutation associated with a given closed path connecting a point $\boldsymbol{R}$ to $\sigma\boldsymbol{R} \cong \boldsymbol{R}$.

[69] For a pedagogical review, see R. MacKenzie and F. Wilczek, *Int. Jour. Mod. Phys. A* **3**, 2827 (1988).

[70] M. G. G. Laidlaw and C. M. DeWitte, *Phys. Rev. D* **3**, 6 (1971)

[71] Y.-S. Wu, *Phys. Rev. Lett.* **52**, 2103 (1984); Y.-S. Wu, **53**, 111 (1984).

In general, if q is a coordinate on a multiply connected space $\mathcal{M}$, the propagator $K(q_2, t_2 \,|\, q_1, t_1)$ may be written as

$$\begin{aligned} K(q_2, t_2 \,|\, q_1, t_1) &= \sum_{\mu \in \pi_1(\mathcal{M})} K_\mu(q_2, t_2 \,|\, q_1, t_1) \\ &= \sum_{\mu \in \pi_1(\mathcal{M})} \chi(\mu) \sum_{q(t) \in \mu} e^{iS[q(t)]/\hbar} \quad , \end{aligned} \tag{3.291}$$

where the sum over μ is over all homotopy sectors in $\pi_1(\mathcal{M})$, and where $\chi(\mu)$ is a unitary one-dimensional representation of $\pi_1(\mathcal{M})$[72]. Thus $\chi(\mu' \circ \mu) = \chi(\mu')\,\chi(\mu)$. In general, $q_1 \neq q_2$ and $q(t)$ is therefore not a closed loop. But by defining a standard set of paths from an arbitrary point $q_0 \in \mathcal{M}$ (assuming $\mathcal{M}$ is connected) to every other point[73], one can append one of these paths or its inverse to the path $q(t)$ to create a closed path. In this way, each path $q(t)$ from q_1 to q_2 can be identified with a homotopy sector.

Paths in configuration space enter the Feynman path integral description of the many-particle propagator, *viz.*

$$K(\boldsymbol{R}', t_2 \,|\, \boldsymbol{R}, t_1) = \frac{1}{N!} \sum_{\sigma \in \mathcal{S}_N} \int_{\boldsymbol{R}}^{\sigma \boldsymbol{R}'} D\boldsymbol{R}(t)\, \exp\left\{ \frac{i}{\hbar} \int_{t_1}^{t_2} dt \left(L(\boldsymbol{R}, \dot{\boldsymbol{R}}, t) + \hbar \frac{\theta}{\pi} \sum_{i<j} \dot{\varphi}_{ij} \right) \right\} \quad , \tag{3.292}$$

where the boundary conditions in the σ sector are given by $\boldsymbol{R}(t_1) = \boldsymbol{R} = \{\boldsymbol{r}_1, \ldots, \boldsymbol{r}_N\}$ and $\boldsymbol{R}(t_2) = \sigma \boldsymbol{R}' = \{\boldsymbol{r}'_{\sigma(1)}, \ldots, \boldsymbol{r}'_{\sigma(N)}\}$. Here, $\varphi_{ij} = \arg(z_i - z_j) = \tan^{-1}[(y_i - y_j)/(x_i - x_j)]$ is the relative angle between particles i and j. The $\dot{\varphi}_{ij}$ term in the Lagrangian keeps track of the relative winding of particles, associating a phase factor $e^{i\theta}$ to each interchange $\Delta\varphi_{ij} = \pi$. Thus, to shift the statistical angle by θ one must alter the many-particle Lagrangian:

$$L(\theta) = L(0) + \hbar \frac{\theta}{\pi} \sum_{i<j} \dot{\varphi}_{ij} \quad . \tag{3.293}$$

Since the additional term is a total time derivative, the angle θ does not appear in the equations of motion. However, the quantity $\dot{\varphi}_{ij}\, dt$ cannot be regarded as an exact differential, because it is not the differential of a *single valued* function of the coordinates $\boldsymbol{R}$. Thus, the 'statistical' part of the action leads to additional phase interference between paths of differing winding number. This is the essence of statistical transmutation.

Charge-flux composites

A compelling realization of fractional statistics was proposed by Wilczek[74] who noted that a composite object consisting of a particle of charge e and a flux tube of strength $\phi = \theta\hbar c/e$ would

[72] Since all one-dimensional representations are abelian, it is really only the abelianized $\pi_1(\mathcal{M})$ which matters here. This is called the *first homology group*, $H_1(\mathcal{M})$.

[73] Such a construction is known as a *standard path mesh*.

[74] F. Wilczek, *Phys. Rev. Lett.* **48**, 1144 (1982); F. Wilczek, *Phys. Rev. Lett.* **49**, 957 (1982).

possess fractional statistics. Recall that when a quantum-mechanical particle of charge q encircles a fixed solenoid of flux ϕ, its wavefunction accrues a phase $e^{iq\phi/\hbar c}$ – this is the celebrated Aharonov-Bohm effect. The same phase would result from a quantum-mechanical solenoid orbiting around a fixed charge. Now consider two of Wilczek's charge-flux composites and compute the phase they generate upon interchange, which is *half* a complete revolution. There are two contributions to the accumulated phase. A factor $e^{ie\phi/2\hbar c} = e^{i\theta/2}$ is generated from the *charge* of particle 1 moving in the field of the *flux* of particle 2, and an identical factor arises from the *flux* of particle 1 moving in the field of the *charge* of particle 2. The net accrued phase is thus $e^{i\theta}$.

A generic Lagrangian $L = \frac{1}{2}m\dot{\boldsymbol{R}}^2 - V(\boldsymbol{R})$, altered to account for fractional statistics as in Eqn. 3.293, results in the many-body Hamiltonian

$$H = \sum_i \frac{1}{2m}\left(\boldsymbol{p}_i - \hbar\,\frac{\theta}{\pi}\sum_j{}' \frac{\hat{\boldsymbol{z}}\times(\boldsymbol{r}_i-\boldsymbol{r}_j)}{|\boldsymbol{r}_i-\boldsymbol{r}_j|^2}\right)^2 + V(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N) \quad , \tag{3.294}$$

where the prime on the sums indicates that the $j=i$ term is to be excluded. The θ-dependent term resembles a 'statistical vector potential'

$$\boldsymbol{\mathcal{A}}_i(\boldsymbol{R}) = \frac{\theta}{\pi}\cdot\frac{\phi_0}{2\pi}\sum_j{}' \frac{\hat{\boldsymbol{z}}\times(\boldsymbol{r}_i-\boldsymbol{r}_j)}{|\boldsymbol{r}_i-\boldsymbol{r}_j|^2} = \frac{\theta}{\pi}\cdot\frac{\hbar c}{e}\sum_j{}' \boldsymbol{\nabla}_i\,\varphi_{ij} \quad , \tag{3.295}$$

where $\phi_0 = hc/e$ is the Dirac flux quantum. The form of the statistical vector potential is the same as the vector potential of a flux tube of strength $\phi = 2\theta\hbar c/e$, which is *twice* the flux of Wilczek's composite. The reason for this is that the statistical vector potential accounts for both the charge-flux and the flux-charge interactions. Note that

$$\boldsymbol{p}_i - \frac{e}{c}\,\boldsymbol{\mathcal{A}}_i(\boldsymbol{R}) = \exp\left(+\,i\,\frac{\theta}{\pi}\sum_j{}'\varphi_{ij}\right)\boldsymbol{p}_i\,\exp\left(-\,i\,\frac{\theta}{\pi}\sum_j{}'\varphi_{ij}\right) \quad , \tag{3.296}$$

indicating that the statistical vector potential is a *pure gauge*, although a *topologically nontrivial* one, because the gauge factor is not single-valued as a function of the coordinates $\boldsymbol{R}$. Application of this singular gauge transformation to a symmetric wavefunction yields a multivalued wavefunction of the kind in Eqn. 3.290.

There are thus two equivalent ways to formulate the implementation of fractional statistics in $d=2$ space dimensions. We can work with single-valued wavefunctions and include a statistical vector potential in our many-body Hamiltonian. This leads to long-ranged two- and (from the $\boldsymbol{\mathcal{A}}^2$ term) three-body interactions. Equivalently, we can employ a singular gauge transformation to 'gauge away' the statistical vector potential at the cost of requiring multivalued wavefunctions, as in Eqn. 3.290. Wilczek named particles obeying fractional statistics *anyons*, presumably because they can have *any* statistics.

Statistical transmutation in field theory

Suppose we have a theory with a conserved current j^μ, which means $\partial_\mu j^\mu = 0$. Here we use the Minkowski metric $\eta^{\mu\nu} = (+,-,-)$ to raise and lower indices, with $x^\mu = (t,x,y)$ and $d^3x = dt\,dx\,dy$. Given a field theory with a conserved matter current j^μ, one can transmute statistics to the matter field by coupling this current to a U(1) gauge field a^μ and adding a Chern-Simons term[75] to the action, *viz.*

$$S_{\rm mat}(\theta) = S_{\rm mat}(0) + \frac{e}{c}\int d^3x\; j^\mu\, a_\mu + \frac{e^2}{4\theta\hbar c^2}\int d^3x\;\epsilon^{\mu\nu\lambda}\,a_\mu\,\partial_\nu\,a_\lambda \quad . \tag{3.297}$$

Although the bare a_μ field is present in the Chern-Simons term, and not only its field strength $f_{\mu\nu} = \partial_\mu a_\nu - \partial_\nu a_\mu$, the action remains gauge-invariant because $\mathcal{J}^\mu = \epsilon^{\mu\nu\lambda}\,a_\mu\partial_\nu a_\lambda$ is a conserved current. Thus, if we make the gauge transformation $a_\mu \to a_\mu + \partial_\mu f$, the change in the Chern-Simons term is

$$S_{\rm CS} \to S_{\rm CS} + \frac{e^2}{4\theta\hbar c^2}\int d^3x\;\epsilon^{\mu\nu\lambda}\,\partial_\mu(f\,\partial_\nu a_\lambda) \quad , \tag{3.298}$$

which vanishes if taken over a closed surface. When taken over a manifold with boundary, an extra contribution must be included at the edge in order to render the action gauge-invariant. We shall discuss this feature further on below.

For example, with nonrelativistic particles we have

$$\begin{aligned} S_{\rm mat}\int d^3x\,\Big\{\sum_{i=1}^N \tfrac{1}{2}m\dot{\boldsymbol{x}}_i^2 - V(\boldsymbol{x}_1,\ldots,\boldsymbol{x}_N)\Big\} \\ j^\mu(x) = \int d\tau\;\sum_{i=1}^N \delta^{(3)}\big(x - x_i(\tau)\big)\,\frac{dx_i^\mu}{d\tau} \quad . \end{aligned} \tag{3.299}$$

Since the action in Eqn. 3.297 is quadratic in the a_μ fields, they can be integrated out simply by solving the equations of motion,

$$\frac{c}{e}\,\frac{\delta S}{\delta a_\mu} = j^\mu + \frac{e}{4\theta\hbar c}\,\epsilon^{\mu\nu\lambda}f_{\nu\lambda} = 0 \quad . \tag{3.300}$$

Thus,

$$j^0 = -\frac{e}{2\theta\hbar c}\,f_{12} = \frac{eb}{2\theta\hbar c} = \frac{\pi}{\theta}\,\frac{b}{\phi_0} \quad , \tag{3.301}$$

where $b = \partial_2 a_1 - \partial_1 a_2 = \partial_x a^y - \partial_y a^x$. One can now integrate out the a_μ fields by manipulating Eqn. 3.300 to obtain

$$\epsilon_{\mu\nu\lambda}\,j^\lambda = \frac{e}{2\theta\hbar c}\,(\partial_\nu a_\mu - \partial_\mu a_\nu) \quad . \tag{3.302}$$

[75] S. Deser, R. Jackiw, and S. Templeton, *Phys. Rev. Lett.* **48**, 975 (1982); J. Schonfeld, *Nucl. Phys.* **B185**, 157 (1981).

Working in the Lorentz gauge $\partial_\mu a^\mu = 0$, we can invert the above relation to yield

$$\Box\, a_\mu = \frac{2\theta\hbar c}{e}\,\eta_{\mu\nu\lambda}\,\partial^\nu j^\lambda \quad , \tag{3.303}$$

where $\Box = c^{-2}\,\partial_t^2 - \boldsymbol{\nabla}^2$ is the wave operator[76]. When substituted into the action, this yields

$$\begin{aligned} S_{\text{eff}}(\theta) &= S_{\text{mat}}(0) + \hbar\,\theta \int d^3x\;\eta_{\mu\nu\lambda}\;j^\mu(x)\;\frac{\partial^\nu}{\Box}\;j^\lambda(x) \\ &= S_{\text{mat}}(0) + 2\hbar\theta N_{\text{link}} \quad , \end{aligned} \tag{3.304}$$

where N_{link} is the *linking number* of the particle trajectories. For a complete revolution of one particle around another, $N_{\text{link}} = 1$, and thus we associate θ with the statistical angle for particle interchange (*i.e.* half a complete revolution).

An explicit calculation is instructive. Define the formally nonlocal operator

$$K^\mu(x - x') = \frac{\partial^\mu}{\Box} \quad , \tag{3.305}$$

which satisfies $\widetilde{\partial}_\nu K^\nu(x - x') = \delta^{(3)}(x - x')$, where $\widetilde{\partial}_\mu = (c^{-2}\,\partial_t\,, \boldsymbol{\nabla})$. Gauge freedom allows us to take $K^0 = 0$ and $K^i(x - x') = k^i(\boldsymbol{x} - \boldsymbol{x}')\,\delta\big(x^0 - x'^0\big)$, with

$$k^i(\boldsymbol{x} - \boldsymbol{x}') = \frac{1}{2\pi}\,\frac{x^i - x'^i}{|\boldsymbol{x} - \boldsymbol{x}'|} \quad . \tag{3.306}$$

The function $k^i(\boldsymbol{x} - \boldsymbol{x}')$ is recognized as the vector potential of a flux tube of unit strength. Now let's wind one particle (X) around another which stays fixed at the origin. The particle currents are then

$$\begin{aligned} j^0(x) &= \delta(\boldsymbol{x}) + \delta(\boldsymbol{x} - \boldsymbol{X}) \\ \boldsymbol{j}(x) &= \delta(\boldsymbol{x} - \boldsymbol{X})\,\dot{\boldsymbol{X}} \quad , \end{aligned} \tag{3.307}$$

where $\boldsymbol{X} = \boldsymbol{X}(\tau)$. The linking number term in Eqn. 3.304 then gives

$$-\int d^3x \int d^3x'\; j^\mu(x)\,\epsilon_{\mu\nu\lambda}\,K^\nu(x - x')\,j^\lambda(x') = -2\int d\tau\;\eta_{ij}\,k^i(-\boldsymbol{X})\,\dot{X}^j = 2\epsilon_{ij}\oint dX^j\,\frac{X^i}{\boldsymbol{X}^2} \tag{3.308}$$

which is indeed $2N_{\text{link}}$. Another way to see it: the geometric flux enclosed by $\boldsymbol{X}(\tau)$ as it winds around the origin is

$$\phi = \oint d\boldsymbol{l}\cdot\boldsymbol{a} = \int dS\;b = \frac{2\theta\hbar c}{e}\int dS\;j^0 = \frac{\nu\theta}{\pi}\,\phi_0 \quad , \tag{3.309}$$

[76] We have to be a bit careful here since $\partial_0 = \partial_t$ and $\partial_{1,2} = \partial_{x,y}$ don't have the same units. The wave operator should be written as $\Box = c^{-2}\,\partial_t^2 - \boldsymbol{\nabla}^2$. The symbol $\eta_{\mu\nu\lambda}$ in Eqn. 3.303 is given by $\eta_{012} = -\eta_{021} = \eta_{120} = -\eta_{210} = 1$ and $\eta_{201} = -\eta_{102} = c^{-2}$, with all other elements vanishing. The simplest way to make everything work out is to pour yourself a nice glass of bourbon and measure space and time in the same units, *i.e.* take $c = 1$.

where the charge is taken to be νe. The Aharonov-Bohm phase is then $e^{2\pi i\nu\phi/\phi_0} = e^{2i\nu^2\theta}$, which is just what we expect.

Another example comes from the $(2+1)$-dimensional $\mathsf{O}(3)$ nonlinear sigma model, with

$$S_{\mathsf{mat}} = \frac{1}{2g}\int d^3x\ (\partial_\mu n^a)(\partial^\mu n^a)$$
$$j^\mu(x) = \frac{1}{8\pi}\,\epsilon_{abc}\,\epsilon^{\mu\nu\lambda}\,n^a\,\partial_\nu n^b\,\partial_\lambda n^c \quad , \tag{3.310}$$

where $\hat{\boldsymbol{n}}(x)$ is a unit vector lying along the surface of a two-dimensional sphere. For this model, which possesses a Lorentz invariance, we define $x^\mu = (ct, \boldsymbol{x})$, $\partial_\mu = \partial/\partial x^\mu$, and $d^3x = c\,dt\,dx\,dy$. The conservation of j^μ licenses us to write

$$j^\mu = \frac{1}{8\pi}\,\epsilon_{abc}\,\epsilon^{\mu\nu\lambda}\,n^a\,\partial_\nu n^b\,\partial_\lambda n^c \equiv \epsilon^{\mu\nu\lambda}\partial_\nu \mathcal{A}_\lambda \quad , \tag{3.311}$$

where $\mathcal{A}_\mu[j]$ is a gauge field. Integrating out the CS gauge field by its equations of motion, we obtain $a_\mu = -(e/2\theta\hbar c)\,\mathcal{A}_\mu$, and hence $S_{\mathsf{eff}}(\theta) = S_{\mathsf{mat}}(0) + \theta S_{\mathsf{Hopf}}$, where

$$S_{\mathsf{Hopf}}/\hbar = -\int d^3x\ j^\mu(x)\,\mathcal{A}_\mu\big[j(x)\big] = -\int d^3x\ \epsilon^{\mu\nu\lambda}\mathcal{A}_\mu\,\partial_\nu\,\mathcal{A}_\lambda \quad , \tag{3.312}$$

which is the so-called *Hopf term*. Note that it is nonlocal in the n^a fields because $\mathcal{A}_\mu[j]$ is a *functional* of its argument. At any fixed time t, we may demand that the field $\hat{\boldsymbol{n}}(\boldsymbol{x}, t)$ approaches the same value $\hat{\boldsymbol{n}}(\infty)$ as $|\boldsymbol{x}| \to \infty$. This compactifies $\mathbb{R}^2 \to S^2$. Now consider the function $\hat{\boldsymbol{n}}(\boldsymbol{x}, t)$ as a function of t for $t \in [0, T]$. Everywhere along this time interval, $\hat{\boldsymbol{n}}(\boldsymbol{x}, t)$ takes its values on the unit sphere S^2 (we assume $\hat{\boldsymbol{n}}(\infty, t) = \hat{\boldsymbol{n}}(\infty)$ is fixed as a function of t). The function $\hat{\boldsymbol{n}}(\boldsymbol{x}, t)$ with t fixed may be regarded as a map taking the compactified real space S^2 to the internal $\hat{\boldsymbol{n}}$ space S^2. The space of such maps is called $\mathcal{Q} = \mathsf{Map}_0(S^2, S^2)$. Now there is a general result which says that[77]

$$\pi_k(\mathcal{Q}) \cong \pi_{k+2}(S^2) \tag{3.313}$$

for all nonnegative integers k. Recall that $\pi_k(\mathcal{M})$ is the group of equivalence classes of maps from $\mathbb{S}^k$ to $\mathcal{M}$. In particular, $\pi_0(\mathcal{Q}) \cong \pi_2(S^2) \cong \mathbb{Z}$, which says that the configuration space of the $(2+1)$-dimensional nonlinear sigma model is disconnected, and separates into individual soliton sectors $\mathcal{Q}_n$ where $n \in \mathbb{Z}$. We also have that $\pi_1(\mathcal{Q}) \cong \pi_3(S^2) \cong \mathbb{Z}$, which says that we can associate a phase $e^{in\theta}$ in the configuration space path integral of Eqn. 3.291 to paths of winding number n. This is precisely what the Hopf term accomplishes.

Many-body theory of the anyon gas

The field theory of the many anyon problem is then given by the Lagrangian density

$$\mathcal{L} = \bar\psi\,(i\hbar D_0 - \tilde\mu)\,\psi - \frac{\hbar^2}{2m}\,\boldsymbol{D}^*\bar\psi\cdot\boldsymbol{D}\psi - v(\bar\psi\psi) + \frac{e^2}{4\theta\hbar c^2}\,\epsilon^{\mu\nu\lambda}\,a_\mu\,\partial_\nu\,a_\lambda - \frac{1}{16\pi}\,F_{\mu\nu}\,F^{\mu\nu} \quad , \tag{3.314}$$

[77] T. R. Govindarajan, R. Shankar, N. Shaji, and M. Sivajumar, *Int. Jour. Mod. Phys. A* **8**, 3965 (1993).

where $D_\mu = \partial_\mu - i\,(e/\hbar c)(A_\mu + a_\mu)$ is the covariant derivative, A_μ is the physical electromagnetic vector potential, $v(\bar\psi\psi)$ is the potential energy, and $\tilde\mu$ is the chemical potential for the ψ field, which may be either fermionic or bosonic. The statistics of the bosons are transmuted to anyons of statistical angle θ (or $\pi+\theta$) by the CS term. With fermionic ψ, the above action serves as a point of departure for the study of the anyon gas. If the physical electromagnetic fields are weak, one may separate the statistical vector potential $a_\mu = a_\mu^{\rm MF} + \delta a_\mu$ into a mean field contribution satisfying $\epsilon^{ji}\,\partial_i\,a_j^{\rm MF} = (\theta/\pi)n\phi_0$, where n is the bulk density, and a fluctuating part δa_μ. One then integrates out the fermion fields ψ and $\bar\psi$, generating an effective action in terms of A_μ and δa_μ. Finally, one may attempt to integrate out the δa_μ fields, generating an effective action in terms of the physical A_μ fields alone, from which one can directly obtain the electromagnetic response functions of anyon gases. These developments are clearly discussed in the article by Fradkin[78]. A particularly interesting conclusion is that the anyon gas for statistical angle $\theta \neq 0, \pi$ should be a superconductor[79]!

To peek just a little bit into how the sausage is made, consider the case of statistical angle $\theta = \pi + \pi/q$ where $q \in \mathbb{Z}$. We now consider Eqn. 3.314 with *fermionic* fields ψ and $\bar\psi$, and with $\theta = \pi/q$. At the mean field level we have $b = n\phi_0/q$ where n is the number density of anyons. In the absence of any external field B, we have a gas of fermions (ψ particles) in a uniform magnetic field b, hence the cyclotron frequency and magnetic length are given by

$$\hbar\omega_{\rm c} = \frac{\hbar e b}{mc} = \frac{2\pi\hbar^2}{m}\,\frac{n}{q} \qquad , \qquad \ell = \sqrt{\frac{\hbar c}{eb}} = \sqrt{\frac{q}{2\pi n}} \quad . \tag{3.315}$$

The LL filling fraction is then $\nu = 2\pi\ell^2 n = q$, which means we have an integer number q of filled LLs. So far, so good.

The ground state energy is then[80]

$$E_0 = N_\phi \sum_{k=0}^{N_\phi - 1} (k + \tfrac{1}{2})\,\hbar\omega_{\rm c} = A\,\frac{\pi\hbar^2}{m}\,n^2 \quad , \tag{3.316}$$

where A is the area and $N_\phi = bA/\phi_0$. Note that the ground state energy per particle is $\varepsilon_0 = E_0/N = \pi\hbar^2/mv$ where $v = 1/n$ is the specific volume (*i.e.* the area per particle). This result is identical to that of a free Fermi gas of the same density. We thus obtain a finite bulk modulus $\mathcal{B}$ and velocity c of first (thermodynamic) sound:

$$\mathcal{B} = v\frac{\partial^2\varepsilon_0}{\partial v^2} = \frac{2\pi\hbar^2}{m}\,n^2 \qquad , \qquad c = \sqrt{\frac{\mathcal{B}}{mn}} = \frac{\hbar}{m}\,\sqrt{2\pi n} \quad . \tag{3.317}$$

Note that these expressions are independent of q and are identical to the corresponding free Fermi gas values. What is missing here is a description of the compressional sound wave;

[78] E. Fradkin, *Phys. Rev. B* **42**, 570 (1990).

[79] A. Fetter, C. Hanna, and R. B. Laughlin, *Phys. Rev. B* **39**, 9679 (1989); Y. Chen, F. Wilczek, E. Witten, and B. I. Halperin, *Int. Jour. Mod. Phys. B* **3**, 1001 (1989).

[80] We assume our anyons are spinless.

our single particle energy spectrum has a gap of $\hbar\omega_{\rm c}$. The sound wave appears in a more sophisticated random phase approximation (RPA) treatment, as first shown by Fetter, Hanna, and Laughlin (1989).

A simple calculation, due to Chen, Wilczek, Witten, and Halperin (1989), shows that despite the breaking of time-reversal symmetry for $\theta = \pi + \pi/q$ (and for all $\theta \neq 0, \pi$), the anyon gas wants to expel magnetic flux. To this end, consider our anyon gas in the presence of an external applied magnetic field B (now parallel to the statistical field b). At the mean field level, the effective field strength is $b + B$, and as B is increased, the LL degeneracy $N_\phi = (b + B)A/\phi_0$ increases. Since the particle number N remains constant, a fraction x of the states in the $q^{\rm th}$ LL, *i.e.* with LL index $q - 1$, will be empty. Number conservation then gives $(q - x)(b + B) = qb$, which determines x. Summing the single particle energies, we obtain

$$E_0(B > 0) = A\,\frac{\pi\hbar^2 n}{m}\left[1 + \frac{B}{qb} - \left(1 - \frac{1}{q}\right)\frac{B^2}{b^2}\right] \tag{3.318}$$

When $B < 0$ and the external field is anti-aligned with b, the effective field strength is $b + B < b$. Now the LL degeneracy N_ϕ is smaller, and number conservation requires that a fraction y of states in the $(q+1)^{\rm th}$ LL (*i.e.* with LL index q) are occupied, with $(n + y)(b + B) = nb$. Summing once again the single-particle energies, one finds

$$E_0(B < 0) = A\,\frac{\pi\hbar^2 n}{m}\left[1 - \frac{B}{qb} - \left(1 + \frac{1}{q}\right)\frac{B^2}{b^2}\right] \quad . \tag{3.319}$$

Thus for general B we have

$$E_0(B) = A\,\frac{\pi\hbar^2 n}{m}\left[1 + \frac{|B|}{qb} - \left(1 - \frac{1}{q}\ \mathsf{sgn}B\right)\frac{B^2}{b^2}\right] \tag{3.320}$$

and any finite B initially increases the total energy. Thus the system always wants to expel a weak external field, despite the fact that time-reversal symmetry is explicitly broken!

To demonstrate the Meissner effect, one must perform substantially more refined calculations. RPA calculations yield a London penetration depth of $\lambda_{\rm L} = \sqrt{mc^2 d/4\pi ne^2}$, where d is the distance between two-dimensional planes, and a sound wave velocity of $c = \frac{\hbar}{m}\sqrt{2\pi n}$, exactly as found above. Thus, the excitation spectrum of the $\theta = \pi + \pi/q$ anyon gas for $q \in \mathbb{Z}_+$ is qualitatively different from the spectrum of a Fermi liquid. While the latter exhibits a gapless sound mode, it also exhibits a continuum of particle-hole excitations extending down to zero energy, and which are responsible for various dissipative processes. In the anyon gas, the particle-hole continuum begins at a finite energy $\hbar\omega_{\rm c} = 2\pi\hbar^2 n/qm$, which properly tends to zero in the fermion limit $q \to \infty$, and the *only* gapless excitation is the density wave. This density wave is a Goldstone mode which in an *uncharged* superconductor (*i.e.* a superfluid) would correspond to phase fluctuations of the order parameter in the phase where $\mathsf{U}(1)$ is

spontaneously broken. When minimally coupled to electromagnetism, this mode is "eaten"[81] via the Anderson-Higgs mechanism, and the photon becomes massive (*i.e.* the Meissner effect).

3.5.3 Cultural interlude

Essentially we have done all the work to derive the CSGL action, which is given in Eqn. 3.314 for a particular choice $\theta = q\pi$ where q is an odd integer and where ψ is a bosonic field. The CS term then transmutes the bosons into fermions, and different choices of odd q, while representing the same theory at the level of $\mathcal{L}$, yield different theories at the mean field level, as we shall soon see. Before engaging with the CSGL theory of the FQHE, though, we take a stroll down memory lane to recall two highlights of the Heroic Era of the FQHE.

Girvin-MacDonald order

It was first suggested by Girvin and MacDonald[82] that Laughlin's wavefunction could be understood as a condensate of composite objects consisting of both charge and flux. Specifically, Girvin and MacDonald showed that if one were to adiabatically pierce each electron in the $\nu = 1/q$ state with a flux tube of strength $q\phi_0$, the resulting off-diagonal density matrix,

$$\begin{aligned}\tilde{n}_1(\boldsymbol{r},\boldsymbol{r}') &= \int d^2r_2 \cdots \int d^2r_N \; \Psi_q^*(\boldsymbol{r},\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) \, \exp\left(-\frac{ie}{\hbar c} \int_{\boldsymbol{r}}^{\boldsymbol{r}'} d\boldsymbol{s}\cdot\boldsymbol{\mathcal{A}}(\boldsymbol{s}) \right) \Psi_q(\boldsymbol{r}',\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) \\ &= \int d^2r_2 \cdots \int d^2r_N \; \tilde{\Psi}_q^*(\boldsymbol{r},\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) \, \tilde{\Psi}_q(\boldsymbol{r}',\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N)\end{aligned} \tag{3.321}$$

where $\boldsymbol{\mathcal{A}}(\boldsymbol{s}) = (q\hbar c/e)\sum_{j=2}^N \boldsymbol{\nabla}\varphi(\boldsymbol{s}-\boldsymbol{r}_j)$ and

$$\tilde{\Psi}_q(\boldsymbol{r},\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) = \exp\left(-iq\sum_{j=2}^N \varphi(\boldsymbol{r}-\boldsymbol{r}_j)\right) \Psi_q(\boldsymbol{r},\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) \quad , \tag{3.322}$$

decays only *algebraically* at long distances, as $|\boldsymbol{r}-\boldsymbol{r}'|^{-q/2}$. This follows from the plasma analogy applied to the product

$$\tilde{\Psi}_q^*(\boldsymbol{r},\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N)\,\tilde{\Psi}_q(\boldsymbol{r}',\boldsymbol{r}_2,\ldots,\boldsymbol{r}_N) \equiv \exp\Big[-\beta\widetilde{H}(\boldsymbol{r},\boldsymbol{r}';\boldsymbol{r}_2,\ldots\boldsymbol{r}_N)\Big] \quad , \tag{3.323}$$

[81] Om nom nom. See E. S. Abers and B. W. Lee, *Phys. Rep.* **9**, 1 (1973).

[82] S. M. Girvin in *The Quantum Hall Effect*, R. Prange and S. M. Girvin, eds. (Springer, 1986); S. M. Girvin and A. H. MacDonald, *Phys. Rev. Lett.* **58**, 1252 (1987).

again with $\beta = 1/m$. One has

$$\widetilde{H}(\boldsymbol{r},\boldsymbol{r}';\boldsymbol{r}_2,\ldots\boldsymbol{r}_N) = -2q^2\sum_{2\le i<j}^{N}\ln|\boldsymbol{r}_i-\boldsymbol{r}_j| + \frac{q}{2\ell^2}\sum_{i=2}^{N}\boldsymbol{r}_i^2 + \frac{q}{4\ell^2}\big(\boldsymbol{r}^2+\boldsymbol{r}'^2\big) - q^2\sum_{i=2}^{N}\Big(\ln|\boldsymbol{r}-\boldsymbol{r}_i| + \ln|\boldsymbol{r}'-\boldsymbol{r}_i|\Big) \quad . \tag{3.324}$$

This corresponds to a system of $(N+1)$ logarithmically interacting charges, and a uniform background with the $(N-1)$ charge $\sqrt{2}\,q$ particles at positions $\{\boldsymbol{r}_2,\ldots\boldsymbol{r}_N\}$, and two charge $\frac{1}{\sqrt{2}}\,q$ particles at positions $\boldsymbol{r}$ and $\boldsymbol{r}'$ *which do not interact with each other*. Adding back this interaction gives the desired result,

$$\tilde{n}_1(\boldsymbol{r},\boldsymbol{r}') = C\,|\boldsymbol{r}-\boldsymbol{r}'|^{-q/2} \quad , \tag{3.325}$$

where C is a dimensionful constant proportional to $\exp(-\beta F)$, where F is the classical free energy of the *fully interacting* $(N+1)$-particle system, *i.e.* where we include the $-2\,(\frac{1}{2}q)^2\ln|\boldsymbol{r}-\boldsymbol{r}'|$ interaction between charge $\frac{1}{\sqrt{2}}\,q$ test particles at $\boldsymbol{r}$ and $\boldsymbol{r}'$. Note that the Girvin-MacDonald result establishes power-law, or *quasi*-long-ranged order. By contrast, the off-diagonal one-body density matrix in the Laughlin state Ψ_q is given by

$$n_1(\boldsymbol{r},\boldsymbol{r}') = \langle\,\Psi_q\,|\,\psi^\dagger(\boldsymbol{r}),\psi(\boldsymbol{r}')\,|\,\Psi_q\,\rangle = \frac{\nu}{2\pi\ell^2}\,G(\boldsymbol{r},\boldsymbol{r}') \quad , \tag{3.326}$$

where $G(\boldsymbol{r},\boldsymbol{r}')$ is given in Eqn. 3.242. This follows from the relation $\langle c_m^\dagger\,c_n\rangle = \nu\,\delta_{m,n}$ in any homogeneous state, where m and n are angular momentum indices. Thus, $n_1(\boldsymbol{r},\boldsymbol{r}')$ falls off as a Gaussian in the Laughlin state Ψ_q, which is very different from the quasi-ODLRO exhibited by $\tilde{n}(\boldsymbol{r},\boldsymbol{r}')$ in the state $\tilde{\Psi}_q$.

Read's order parameter

Nicholas Read[83] proposed the order parameter operator

$$\phi^\dagger(\boldsymbol{r}) = \psi^\dagger(\boldsymbol{r})\,U^q(z) \tag{3.327}$$

for the Laughlin state Ψ_q, where $\psi^\dagger(\boldsymbol{r})$ is an electron creation operator in second quantized form and $U(z) = \prod_{i=1}^{N}(z-z_i)$ is the quasihole creation operator in first quantized form. Defining the $(N+1)$-particle state,

$$|\,\Phi_q[N+1]\,\rangle = \int d^2r\;e^{-r^2/4\ell^2}\;\psi^\dagger(\boldsymbol{r})\,U^q(z)\,|\,\Psi_q[N]\,\rangle \quad , \tag{3.328}$$

[83] N. Read, *Phys. Rev. Lett.* **62**, 86 (1989).

Now let's calculate the overlap $\Phi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N+1}) = \langle\, \boldsymbol{r}_1,\ldots \boldsymbol{r}_{N+1} \,|\, \Psi_q[N+1]\,\rangle$. We note that

$$\begin{aligned}\psi(\boldsymbol{r})\,|\,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N+1}\,\rangle &= \frac{1}{\sqrt{(N+1)!}}\,\psi(\boldsymbol{r})\,\psi^\dagger(\boldsymbol{r}_{N+1})\cdots\psi^\dagger(\boldsymbol{r}_1)\,|\,0\,\rangle \\ &= \frac{1}{2\pi\ell^2\sqrt{(N+1)}}\Big\{G(\boldsymbol{r},\boldsymbol{r}_{N+1})\,|\,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N\,\rangle \\ &\qquad - G(\boldsymbol{r},\boldsymbol{r}_N)\,|\,\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N-1},\boldsymbol{r}_N\,\rangle + \ldots + (-1)^N\,G(\boldsymbol{r},\boldsymbol{r}_1)\,|\,\boldsymbol{r}_2,\ldots,\boldsymbol{r}_{N+1}\,\rangle\Big\}\quad,\end{aligned} \tag{3.329}$$

where $G(\boldsymbol{r},\boldsymbol{r}') = \big\{\psi(\boldsymbol{r}),\psi^\dagger(\boldsymbol{r}')\big\}$ is given in Eqn. 3.242. Recall that $G(\boldsymbol{r},\boldsymbol{r}')$ is the Girvin-Jach reproducing kernel for analytic functions, *viz.*

$$\int d^2r'\,G(\boldsymbol{r},\boldsymbol{r}')\,f(z')\,e^{-|z'|^2/4\ell^2} = f(z)\,e^{-|z|^2/4\ell^2}\quad. \tag{3.330}$$

Thus,

$$\begin{aligned}\Phi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N+1}) &= \frac{1}{\sqrt{(N+1)}}\int d^2r\,e^{-\boldsymbol{r}^2/4\ell^2}\Big\{G(\boldsymbol{r},\boldsymbol{r}_{N+1})\,\Psi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)\prod_{\substack{i=1\\(i\neq N+1)}}^{N+1}(z-z_i)^q \\ &\qquad - G(\boldsymbol{r},\boldsymbol{r}_N)\,\Psi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N-1},\boldsymbol{r}_{N+1})\prod_{\substack{i=1\\(i\neq N)}}^{N+1}(z-z_i)^q + \ldots\Big\} \\ &= (N+1)^{-1/2}\left\{\prod_{\substack{i=1\\(i\neq N+1)}}^{N+1}(z_{N+1}-z_i)^q\,\Psi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_N)\,e^{-\boldsymbol{r}_{N+1}^2/4\ell^2}\right. \\ &\qquad \left. - \prod_{\substack{i=1\\(i\neq N)}}^{N+1}(z_N-z_i)^q\,\Psi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N-1},\boldsymbol{r}_{N+1})\,e^{-\boldsymbol{r}_N^2/4\ell^2}+\ldots\right\} \\ &= (N+1)^{1/2}\,\Psi_q(\boldsymbol{r}_1,\ldots,\boldsymbol{r}_{N+1})\quad.\end{aligned} \tag{3.331}$$

This says that the Laughlin state may be written as

$$|\,\Psi_q[N]\,\rangle = \frac{1}{\sqrt{N!}}\left(\int d^2r\,e^{-\boldsymbol{r}^2/4\ell^2}\,\psi^\dagger(\boldsymbol{r})\,U^q(z)\right)^N|\,0\,\rangle\quad, \tag{3.332}$$

which is a condensate of the composite $\phi^\dagger(\boldsymbol{r})$, which creates q quasiholes and fills them with one electron, thus increasing the electron number by 1. What makes Read's operator so natural is that it is truly a boson and thus can condense. Let's compute the statistical angle Θ resulting from the interchange of two Read composites. There are four contributions:

$$\Theta = \pi + q\pi + q\pi + q^2\theta = (q+1)\,\pi \bmod 2\pi\quad. \tag{3.333}$$

The first of these contributions arises from the exchange of the fermions created by the electron creation operators $\psi^\dagger$. Next, there is a phase accrued by the fermions moving in the field of the flux tubes effectively added by the U^q operators. A single electron encircling q Dirac flux tubes accrues a phase angle of $2\pi q$, and so an exchange gives us half this value, or $q\pi$. However, there is an equal phase arising from the flux of one composite orbiting the charge of the other. Finaly, there is the statistical angle due to the exchange of the quasiparticles themselves. Since each composite consists of an electron plus q quasiholes, this last contribution to the statistical angle is $q^2(\pi/q) = q\pi$ since $\theta = \pi/q$. The net statistical angle is thus $\Theta = (q+1)\pi$, which is bosonic when q is odd[84].

3.5.4 The CSGL action

Read's order parameter $\phi^\dagger(\boldsymbol{r}) = \psi^\dagger(\boldsymbol{r})\, U^q(z)$ describes an electron bound to q flux quanta. We may now define a fictitious gauge field $\boldsymbol{a}(\boldsymbol{r})$ whose curl, $b = \hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\boldsymbol{a}$, satisfies $\boldsymbol{\nabla}\times\boldsymbol{a} = q\phi_0\, n(\boldsymbol{r})\hat{\boldsymbol{z}}$, hence

$$b(\boldsymbol{r}) = q\phi_0 \sum_{i=1}^{N} \delta(\boldsymbol{r} - \boldsymbol{r}_i)\, \hat{\boldsymbol{z}} \quad . \tag{3.334}$$

We thus arrive at the cartoon sketched in Fig. 3.12. Suppose $\dot{N}$ is the number current of charge-flux composites moving across the surface Σ. From Faraday's law $\boldsymbol{\nabla}\times\boldsymbol{E} = -\frac{1}{c}\frac{\partial \boldsymbol{B}}{\partial t}$, we have

$$\int\limits_{\Sigma} d\boldsymbol{l}\cdot\boldsymbol{E} = -\frac{1}{c}\,\frac{\partial\Phi}{\partial t} = -\frac{\phi}{c}\,\dot{N} = V_{\mathsf{H}} \tag{3.335}$$

with $\phi = q\phi_0$. We conclude that the Hall conductance is quantized:

$$G_{\mathsf{H}} = \frac{I_{\mathrm{c}}}{V_{\mathsf{H}}} = \frac{-e\dot{N}}{-q(h/e)\dot{N}} = \frac{e^2}{qh} \quad . \tag{3.336}$$

The binding of flux to charge also qualitatively explains why the FQH system is *incompressible*: if the wavefunction Ψ condenses into a superconducting state, *excess* 'magnetic' flux is expelled. But the magnetic flux is proportional to the particle density, hence there must be a fixed uniform density of particles.

The CSGL action functional is given by

$$S_{\mathsf{CSGL}}[\Psi,\Psi^*,A^\mu,a^\mu] = S_{\mathsf{mat}}[\Psi,\Psi^*,A^\mu,a^\mu] + S_{\mathsf{CS}}[a^\mu] \quad , \tag{3.337}$$

[84]Note that the generalization of Read's order parameter to the even denominator bosonic FQHE gives $\Theta = q\pi$ with q even, because the field operator $\psi^\dagger$ then creates a boson. So again the order parameter is bosonic.

where

$$S_{\rm mat}[\Psi,\Psi^*,A^\mu,a^\mu] = \int\!dt\!\int\!d^2x\,\left\{\Psi^*\big(i\hbar\,\partial_t + \tfrac{e}{c}\,A^0 + \tfrac{e}{c}\,a^0\big)\,\Psi - \frac{1}{2m^*}\Big|\big(\tfrac{\hbar}{i}\,\boldsymbol{\nabla} + \tfrac{e}{c}\boldsymbol{A} + \tfrac{e}{c}\,\boldsymbol{a}\big)\,\Psi\Big|^2\right\}$$
$$- \frac{1}{2}\int\!dt\!\int\!d^2x\!\int\!d^2x'\,\big(|\Psi(\boldsymbol{x},t)|^2 - n_0\big)\,v(\boldsymbol{x}-\boldsymbol{x}')\,\big(|\Psi(\boldsymbol{x}',t)|^2 - n_0\big) \quad (3.338)$$

and

$$S_{\rm CS}[a^\mu] = \frac{\pi e}{2\theta c\phi_0}\int\!dt\int\!d^2x\,\epsilon^{\mu\nu\lambda}\,a_\mu\partial_\nu a_\lambda \quad . \quad (3.339)$$

The CS term transmutes the statistical angle of the Ψ field from bosonic to the value θ, which we may choose as suits our nefarious purposes. For fermionic statistics, we require $\theta = q\pi$ with q an odd integer. The background number density is n_0, and the terms linearly proportional to n_0 determine the chemical potential $\mu = n_0\!\int\!d^2x\,v(\boldsymbol{x}) = \hat{v}(\boldsymbol{0})\,n_0$; in Coulomb systems where $\hat{v}(\boldsymbol{0})$ diverges, the average number density $\langle\,|\Psi|^2\,\rangle$ must be n_0 in order to enforce global charge neutrality. Note that the variation of the CS term is given by

$$\delta S_{\rm CS}[a^\mu] = \frac{\pi e}{\theta c\phi_0}\int\!dt\int\!d^2x\,\Big(-b\,\delta a^0 + c\,e^y\,\delta a^x - c\,e^x\,\delta a^y\Big) \quad , \quad (3.340)$$

where

$$\boldsymbol{e} = -\frac{1}{c}\left(\boldsymbol{\nabla}a^0 + \frac{\partial\boldsymbol{a}}{\partial t}\right) \qquad , \qquad b = \hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\boldsymbol{a} = \frac{\partial a^y}{\partial x} - \frac{\partial a^x}{\partial y} \quad (3.341)$$

are the fictitious electric and magnetic field strengths associated with the CS gauge field. Note that we have chosen our factors such that $\boldsymbol{e}$ and b have the same dimensions as the physical electromagnetic field strengths $\boldsymbol{E}$ and B (in cgs units[85]).

The bosonic number density and number current are given by

$$n(\boldsymbol{x},t) = +\frac{c}{e}\,\frac{\delta S}{\delta A^0(\boldsymbol{x},t)} = |\Psi(\boldsymbol{x},t)|^2$$
$$\boldsymbol{j}(\boldsymbol{x},t) = -\frac{c}{e}\,\frac{\delta S}{\delta\boldsymbol{A}(\boldsymbol{x},t)} = \frac{\hbar}{m^*}\,{\rm Im}\big[\Psi^*(\boldsymbol{x},t)\,\boldsymbol{\nabla}\,\Psi(\boldsymbol{x},t)\big] + \frac{e}{m^*c}\,|\Psi(\boldsymbol{x},t)|^2\,\big(\boldsymbol{A}(\boldsymbol{x},t) + \boldsymbol{a}(\boldsymbol{x},t)\big) \quad (3.342)$$

Thus the functional variation of the complete CSGL action with respect to the components of the gauge field a^μ is given by

$$\frac{c}{e}\,\frac{\delta S}{\delta a^0(\boldsymbol{x},t)} = n(\boldsymbol{x},t) - \frac{\pi}{\theta}\,\frac{b(\boldsymbol{x},t)}{\phi_0}$$
$$-\frac{c}{e}\,\frac{\delta S}{\delta\boldsymbol{a}(\boldsymbol{x},t)} = \boldsymbol{j}(\boldsymbol{x},t) - \frac{\pi}{\theta}\,\frac{c\,\boldsymbol{e}(\boldsymbol{x},t)\times\hat{\boldsymbol{z}}}{\phi_0} \quad , \quad (3.343)$$

[85] Which are God's units.

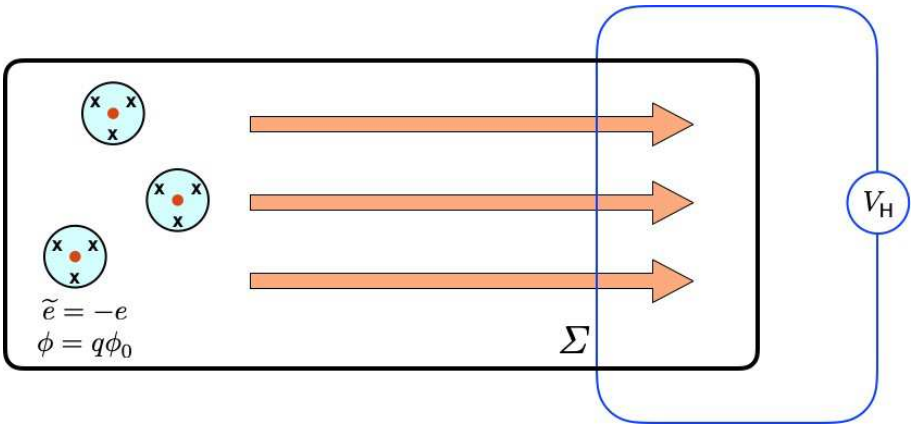


Figure 3.12: Schematic picture of transport for charge-flux composites. The charge current across the surface Σ is $I_c = -e\,\partial_t N$ while the vortex current is $I_v = \phi\,\partial_t N$.

We may also vary the action with respect to the field $\Psi^*(\boldsymbol{x})$:

$$\frac{\delta S}{\delta \Psi^*(\boldsymbol{x},t)} = \left[i\hbar\frac{\partial}{\partial t} + \frac{e}{c}\,(A^0 + a^0) - \frac{1}{2m^*}\left(\frac{\hbar}{i}\,\boldsymbol{\nabla} + \frac{e}{c}\,(\boldsymbol{A}+\boldsymbol{a})\right)^2 \right]\Psi(\boldsymbol{x},t) \\ - \left[\int d^2x'\, v(\boldsymbol{x}-\boldsymbol{x}')\big(|\Psi(\boldsymbol{x}',t)|^2 - n_0\big)\right]\Psi(\boldsymbol{x},t) \quad . \tag{3.344}$$

3.5.5 Mean field solution

We now write $a_\mu = a_\mu^{\rm MF} + \delta a_\mu$ and $\Psi = \Psi^{\rm MF} + \delta\Psi$. The mean field level solution is obtained by setting $\Psi^{\rm MF} = \sqrt{n_0}\,e^{i\phi}$ where ϕ is an arbitrary phase, and setting $A_\mu + a_\mu^{\rm MF} = 0$. In the Ginzburg-Landau theory of superconductivity, ϕ would be the phase of the superconducting condensate. We are free to choose a gauge in which $A_0 = a_0^{\rm MF} = 0$, but for the spatial components the condition $\boldsymbol{A} + \boldsymbol{a}^{\rm MF} = 0$ requires that $B + b^{\rm MF} = 0$, *i.e.* the Chern-Simons magnetic field b *cancels* the applied magnetic field $-B$. Note that we are again taking the applied field as $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$, whereas $\boldsymbol{b} = +b\,\hat{\boldsymbol{z}}$. So $b^{\rm MF} = B$. But we also have the condition from the first of Eqns. 3.343, which demands that

$$b = B = \frac{\theta}{\pi}\,n_0\,\phi_0 = q n_0\,\phi_0 \quad , \tag{3.345}$$

where $\theta = \pi q$ for Fermi statistics, with q odd. Note that in the original field theory, the CS term attaches an infinitely narrow flux tube to each point particle. Therefore the Aharonov-Bohm flux is only sensitive to $\theta \bmod 2\pi$, and all odd integer q are equivalent. Not so at the mean field level! Each odd q results in a *different* mean field solution, with $N_\phi = BA/\phi_0 = qN$, *i.e.* $\nu = N/N_\phi = q^{-1}$.

Suppose the condensate field $\Psi(\boldsymbol{x})$ contains a vortex centered at the origin. Asymptotically, as

$r = |\boldsymbol{x}| \to \infty$, we must then have $\Psi(r,\varphi) = \sqrt{n_0}\,\exp(\pm i\varphi)$ and $\delta\boldsymbol{a}(r,\varphi) = \pm(\hbar c/e)\,\hat{\boldsymbol{\varphi}}/r$, where φ is the azimuthal angle. But this corresponds to a variation in the total CS flux $\Phi_{\rm cs}$, with

$$\delta\Phi_{\rm cs} = \delta\int d^2x\, b(\boldsymbol{x}) = \oint\limits_{r=\infty} d\boldsymbol{l}\cdot\delta\boldsymbol{a} = \phi_0 \quad . \tag{3.346}$$

But then the first of Eqns. 3.343 says that there is a concomitant change in particle number, with

$$\delta N = \frac{\pi}{\theta}\,\frac{\delta\Phi_{\rm cs}}{\phi_0} = \frac{1}{q} \quad . \tag{3.347}$$

This is the fractionally charged quasiparticle! The antivortex corresponds to the quasihole.

3.5.6 Fluctuations about the mean field

Let us write $A^\mu = \bar{A}^\mu + \delta A^\mu$ where $\bar{A}^\mu$ is the electromagnetic 3-vector potential corresponding to uniform magnetic field $\boldsymbol{B} = -B\hat{\boldsymbol{z}}$ and electric field $\boldsymbol{E} = 0$, with B fixed by the condition $\nu = n\phi_0/B = 1/q$ with q an odd integer. Similarly we write $a^\mu = \bar{a}^\mu + \delta a^\mu$ with $\bar{a}_\mu = -\bar{A}_\mu = a^{\rm MF}_\mu$. The action is then

$$\begin{aligned} S = \int d^3x\,\Big\{\hbar\,\Psi^*\big(i\partial_t + \tfrac{e}{\hbar c}\,\delta\mathsf{a}_0\big)\Psi - \frac{\hbar^2}{2m}\,\Big|\big(i\boldsymbol{\nabla} - \tfrac{e}{\hbar c}\,\delta\vec{\mathsf{a}}\big)\Psi\Big|^2\Big\} \\ - \frac{1}{2}\int dt\int d^2x\int d^2x'\,\big(|\Psi(\boldsymbol{x},t)|^2 - n_0\big)\,v(\boldsymbol{x}-\boldsymbol{x}')\,\big(|\Psi(\boldsymbol{x}',t)|^2 - n_0\big) \\ + \frac{\pi e}{2\theta c\phi_0}\int d^3x\,\epsilon^{\mu\nu\lambda}\Big[(\delta\mathsf{a}_\mu - \delta A_\mu)\,\partial_\nu\,(\delta\mathsf{a}_\lambda - \delta A_\lambda) - 2\,\delta\mathsf{a}_\mu\,\partial_\nu\,\bar{A}_\lambda\Big] \quad , \end{aligned} \tag{3.348}$$

where $\delta\mathsf{a}_\mu \equiv \delta a_\mu + \delta A_\mu$. We now follow the method outlined in §3.5.1, now in real rather than Euclidean time, writing $\Psi = \sqrt{n}\,e^{i\phi}$, which results in the matter component of the action

$$\begin{aligned} S_{\rm mat} = \int d^3x\,\Big\{ - \hbar n\big(\partial_t\phi - \tfrac{e}{\hbar c}\,\delta\mathsf{a}_0\big) - \frac{\hbar^2 n}{2m^*}\,\big(\boldsymbol{\nabla}\phi + \tfrac{e}{\hbar c}\,\delta\vec{\mathsf{a}}\big)^2 - \frac{\hbar^2}{8m^* n}\,(\boldsymbol{\nabla}n)^2\Big\} \\ - \frac{1}{2}\int dt\int d^2x\int d^2x'\,\big(n(\boldsymbol{x},t) - n_0\big)\,v(\boldsymbol{x}-\boldsymbol{x}')\,\big(n(\boldsymbol{x}',t) - n_0\big) \quad . \end{aligned} \tag{3.349}$$

We again use a Hubbard-Stratonovich transformation to replace

$$-\frac{\hbar^2 n}{2m^*}\,\big(\boldsymbol{\nabla}\phi + \tfrac{e}{\hbar c}\,\delta\vec{\mathsf{a}}\big)^2 \quad\longrightarrow\quad \frac{m^*\boldsymbol{Q}^2}{2n} - \hbar\boldsymbol{Q}\cdot\big(\boldsymbol{\nabla}\phi + \tfrac{e}{\hbar c}\,\delta\vec{\mathsf{a}}\big) \tag{3.350}$$

Integrating by parts, the contribution from $\phi_{\rm sw}$ in the Lagrangian density is $\hbar(\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{Q})\,\phi_{\rm v}$, hence integrating out $\phi_{\rm v}$ results in the constraint $\partial_t n + \boldsymbol{\nabla}\cdot\boldsymbol{Q} = 0$, which we solve by writing

$$K^\mu \equiv (n,\boldsymbol{Q}) \equiv n_0\,\epsilon^{\mu\nu\lambda}\,\partial_\nu\,(\mathcal{W}_\lambda + w_\lambda) \quad . \tag{3.351}$$

Here n, Q_x, and Q_y are the three components of the 3-vector K^μ.[86] Here $w^\mu = (0, 0, -x)$ is a background vector potential corresponding to electric field $\boldsymbol{e} = 0$ and magnetic field $b = -1$ which will help us keep track of the Berry phase as a vortex in the Ψ field moves in the presence of the background condensate, as in §3.5.1. We may write

$$\begin{aligned} n - n_0 &= n_0 \, (\partial_x \mathcal{W}_y - \partial_y \mathcal{W}_x) \equiv -n_0 \, \mathcal{B} \\ Q_x &= n_0 \, (\partial_y \mathcal{W}_t - \partial_t \mathcal{W}_y) \equiv -n_0 \, u \, \mathcal{E}_y \\ Q_y &= n_0 \, (\partial_t \mathcal{W}_x - \partial_x \mathcal{W}_t) \equiv +n_0 \, u \, \mathcal{E}_x \quad , \end{aligned} \tag{3.352}$$

where u, to be determined, has dimensions of speed. The dimensionless quantities $\boldsymbol{\mathcal{E}}$ and $\mathcal{B}$ are the "electric" and "magnetic" fields derived from the vector potential $\mathcal{W}^\mu$.

Again we separate $\phi = \phi_{\mathsf{sw}} + \phi_{\mathsf{v}}$ into a smooth 'spin-wave' part and a singular vortex part, where

$$J^\mu_{\mathsf{v}} = \frac{1}{2\pi} \, \epsilon^{\mu\nu\lambda} \, \partial_\nu \, \partial_\lambda \, \phi_{\mathsf{v}} = \sum_i q_i \left\{ \begin{array}{c} 1 \\ \dot{\boldsymbol{X}}_i \end{array} \right\} \delta\big(\boldsymbol{x} - \boldsymbol{X}_i(\tau)\big) \quad . \tag{3.353}$$

The matter component of the action may now be written

$$\begin{aligned} S_{\mathsf{mat}} = \int d^3x \, \Bigg\{ & \frac{e n_0}{c} \, \epsilon^{\mu\nu\lambda} \, \delta \mathsf{a}_\mu \, \partial_\nu \, (\mathcal{W}_\lambda + w_\lambda) + \tfrac{1}{2} \, n_0 \, m^* u^2 \, \frac{\boldsymbol{\mathcal{E}}^2}{1 - \mathcal{B}} - \frac{\hbar^2 n_0}{8 m^*} \, \frac{(\boldsymbol{\nabla} \mathcal{B})^2}{1 - \mathcal{B}} \\ & - 2\pi\hbar n_0 \, J^\mu_{\mathsf{v}} \, (\mathcal{W}_\mu + w_\mu) \Bigg\} - \tfrac{1}{2} n_0^2 \int dt \int d^2x \int d^2x' \; \mathcal{B}(\boldsymbol{x}, t) \, v(\boldsymbol{x} - \boldsymbol{x}') \, \mathcal{B}(\boldsymbol{x}', t) \quad . \end{aligned} \tag{3.354}$$

If the Fourier transform of the potential $\hat{v}(\boldsymbol{k})$ is finite at $\boldsymbol{k} = 0$, the constant u may be defined by $u \equiv \sqrt{n_0 \, \hat{v}(\boldsymbol{0})/m^*}$, as in §3.7. The term proportional to $(\boldsymbol{\nabla}\mathcal{B})^2$ is fourth order in derivatives of $\mathcal{W}^\mu$; this term of course yields the crossover to a ballistic dispersion in the phonon spectrum of the superfluid in the ultraviolet regime. In the infrared, we may neglect it, and we have

$$\begin{aligned} S_{\mathsf{mat}} &= \int d^3x \, \Big\{ \tfrac{1}{2} \, n_0 \, m^* u^2 \big(\boldsymbol{\mathcal{E}}^2 - \mathcal{B}^2\big) + \frac{e n_0}{c} \, \epsilon^{\mu\nu\lambda} \, \delta \mathsf{a}_\mu \, \partial_\nu \, (\mathcal{W}_\lambda + w_\lambda) - 2\pi\hbar n_0 \, J^\mu_{\mathsf{v}} \, (\mathcal{W}_\mu + w_\mu) \Big\} \\ S_{\mathsf{CS}} &= \frac{\pi e}{2\theta c \phi_0} \int d^3x \, \epsilon^{\mu\nu\lambda} \Big\{ (\delta \mathsf{a}_\mu - \delta A_\mu) \, \partial_\nu \, (\delta \mathsf{a}_\lambda - \delta A_\lambda) - 2 \, \delta \mathsf{a}_\mu \, \partial_\nu \, \bar{A}_\lambda \Big\} \end{aligned} \tag{3.355}$$

We now integrate out the gauge field $\delta \mathsf{a}_\mu$ using the equations of motion, which yield

$$\begin{aligned} \delta a_\mu &= \delta A_\mu - \frac{\theta}{\pi} \, n_0 \, \phi_0 \, (\mathcal{W}_\mu + w_\mu) + \bar{A}_\mu \\ &= \delta A_\mu - \frac{\theta}{\pi} \, n_0 \, \phi_0 \, \mathcal{W}_\mu \quad , \end{aligned} \tag{3.356}$$

[86] On spatial 2-vectors like $\boldsymbol{Q}$ and the electric field $\boldsymbol{E}$, we do not distinguish between raised and lowered indices, hence $Q^x = Q_x$ *etc.* We shall endeavor to always use lowered indices for such 2-vectors.

since $B = \frac{\theta}{\pi}\, n_0\, \phi_0$ yields $\bar{A}_\mu = \frac{\theta}{\pi}\, n_0\, \phi_0\, w_\mu$. Substituting this back into the action, we obtain the Lagrangian density

$$\mathcal{L}_{\rm eff} = \tfrac{1}{2}\, n_0\, m^* u^2 \big(\boldsymbol{\mathcal{E}}^2 - \mathcal{B}^2\big) - \hbar\theta\, n_0^2\, \epsilon^{\mu\nu\lambda}\, \mathcal{W}_\mu \partial_\nu \mathcal{W}_\lambda + \frac{e n_0}{c}\, \epsilon^{\mu\nu\lambda}\, \delta A_\mu \partial_\nu \mathcal{W}_\lambda - 2\pi\hbar n_0\, J^\mu_{\rm v}\, (\mathcal{W}_\mu + w_\mu) \quad . \tag{3.357}$$

Note that there are two quantities here with dimensions of speed: c and u, with $u \ll c$.

If $\hat{v}(\mathbf{0})$ diverges, which is indeed the case when $v(\boldsymbol{r}) = e^2/\epsilon r$ and $\hat{v}(\boldsymbol{k}) = 2\pi e^2/\epsilon k$, then there is no effective Lorentz symmetry in the superfluid component of the matter Lagrangian. In this case we have

$$S_{\rm eff} = \int d^3x \left\{ \tfrac{1}{2}\, n_0\, m^* \big(\boldsymbol{\nabla}\mathcal{W}^0 + \partial_t \boldsymbol{\mathcal{W}}\big)^2 - \hbar\theta\, n_0^2\, \epsilon^{\mu\nu\lambda}\, \mathcal{W}_\mu \partial_\nu \mathcal{W}_\lambda + \frac{e n_0}{c}\, \epsilon^{\mu\nu\lambda}\, \delta A_\mu \partial_\nu \mathcal{W}_\lambda \right.$$
$$\left. - 2\pi\hbar n_0\, J^\mu_{\rm v}\, (\mathcal{W}_\mu + w_\mu) \right\} - \tfrac{1}{2} n_0^2 \int dt \int \frac{d^2q}{(2\pi)^2}\, \hat{v}(\boldsymbol{q}) \left| \boldsymbol{q} \times \hat{\boldsymbol{\mathcal{W}}}(\boldsymbol{q}) \right|^2 \quad . \tag{3.358}$$

At this point, we may integrate out the gauge field $\mathcal{W}_\mu$. Since the Maxwell term in Eqn. 3.357 or its corrected version in Eqn. 3.358 both involve one higher derivative than the induced Chern-Simons term $\mathcal{W}d\mathcal{W}$, we will ignore the former. Varying with respect to $\mathcal{W}_\mu$, we obtain

$$\epsilon^{\mu\nu\lambda}\, \partial_\nu \mathcal{W}_\lambda = -\frac{\pi}{\theta n_0} \Big(J^\mu_{\rm v} - \frac{1}{2\phi_0}\, \epsilon^{\mu\nu\lambda}\, \delta F_{\nu\lambda} \Big) \quad , \tag{3.359}$$

where $\delta F_{\nu\lambda} = \partial_\nu\, \delta A_\lambda - \partial_\lambda\, \delta A_\nu$ is the field strength tensor corresponding to δA_μ. Thus

$$\mathcal{W}_\mu = \frac{\pi}{\theta n_0} \left(\frac{1}{\phi_0}\, \delta A_\mu - \eta_{\mu\nu\lambda}\, \frac{\partial^\nu}{\Box}\, J^\lambda_{\rm v} \right) \quad . \tag{3.360}$$

Inserting this into $\mathcal{L}_{\rm eff}$ yields our final result[87],

$$\mathcal{L}(J_{\rm v}, \delta A) = -2\pi\hbar n_0\, J^\mu_{\rm v}\, w_\mu - \frac{\pi}{\theta}\, \frac{e}{c}\, J^\mu_{\rm v}\, \delta A_\mu + \frac{e^2}{4\theta\hbar c^2}\, \epsilon^{\mu\nu\lambda}\, \delta A_\mu\, \partial_\nu\, \delta A_\lambda + \hbar\, \frac{\pi^2}{\theta}\, \eta_{\mu\nu\lambda}\, J^\mu_{\rm v}\, \frac{\partial^\nu}{\Box}\, J^\lambda_{\rm v} \quad . \tag{3.361}$$

Thus, we find:

- The vortices of the CSGL condensate accrue a Berry phase in traversing a loop $\mathcal{C}$ of $\gamma_{\mathcal{C}} = -2\pi\, N_0(\mathcal{C})$, where $N_0(\mathcal{C}) = n_0\, A(\mathcal{C})$ is the condensate number density times the area enclosed by the loop, which is to say the average number of condensate particles encircled.
- The vortex current $J^\mu_{\rm v}$ is minimally coupled to fluctuations δA_μ in the physical electromagnetic field, with an effective charge $e^* = \pi e/\theta = e/q$ in the Laughlin state at $\nu = 1/q$.

[87] Recall that $\Box = c^{-2}\, \partial_t^2 - \boldsymbol{\nabla}^2$ and $\eta_{012} = -\eta_{021} = \eta_{120} = -\eta_{210} = 1$ with $\eta_{201} = -\eta_{102} = c^{-2}$.

- There is an induced Chern-Simons term in the physical electromagnetic vector potential. Varying with respect to the physical electromagnetic field, we have[88]

$$J^\mu = -c\,\frac{\delta S}{\delta A_\mu} = -\frac{\pi}{\theta}\,eJ^\mu_{\rm v} + \frac{e^2}{2\hbar c}\,\epsilon^{\mu\nu\lambda}\,\partial_\mu\,\delta A_\lambda \quad , \tag{3.362}$$

  and with the physical electric field given by $\boldsymbol{E} = -c^{-1}(\boldsymbol{\nabla}\delta A^0 + \partial_0\,\delta\boldsymbol{A})$, we obtain

$$\boldsymbol{J} = \frac{e^2}{qh}\,\hat{\boldsymbol{z}}\times\boldsymbol{E} - \frac{e}{q}\,\boldsymbol{J}_{\rm v} \quad . \tag{3.363}$$

  This tells us that in the absence of vortices (*i.e.* quasiparticles) there is a quantized Hall effect $\sigma_{yx} = e^2/qh$. Since vortices are charged, when present they carry an electrical current. However, the random potential from the displaced dopant ions produces many *pinning sites* for vortices, and a pinned vortex, which remains spatially localized, carries zero current.

- Comparing with Eqn. 3.304, we see that the vortices of our theory are anyons with a statistical angle $\vartheta = \pi^2/\theta$. Thus for $\theta = q\pi$ we have $\vartheta = \pi/q$, exactly as the adiabatic calculation of §3.3.6 concluded. Had we included the Maxwell term, the combination of CS and Maxwell terms would have attached a smeared flux tube to each of the vortices, where the length scale of the smearing is $d = m^*u^2/\theta\hbar n_0 c$.

What we *don't* have here is the $1/r$ Coulomb interaction between the vortices, which as we have seen possess finite charge $\pm e/q$. In fact, this is indeed included in the last term of the action of Eqn. 3.355, but we have neglected the long range part of $v(\boldsymbol{r}-\boldsymbol{r}')$ in deriving Eqn. 3.361. More on this below in §3.5.8.

Remarks on units

In our units, $x^\mu = (t, x, y)$ and $d^3x = dt\,dx\,dy$. The units for the components of particle 3-current and vector potentials are

$$[n] = [j^0] = L^{-2} \quad , \quad [\boldsymbol{j}] = L^{-1}T^{-1} \quad , \quad [eA^0] = ELT^{-1} \quad , \quad [e\boldsymbol{A}] = E \quad , \tag{3.364}$$

where L stands for length, T for time, and E for energy. Thus $\left[\frac{e}{c}\,j_\mu A^\mu\right] = EL^{-2}$, *i.e.*energy density. Since $\alpha^{-1} = \hbar c/e^2 \approx 137.036$ is the inverse fine structure constant, we have

$$[e^2] = [\hbar c] = EL \quad \Rightarrow \quad [e] = E^{1/2}L^{1/2} \quad . \tag{3.365}$$

Note that the physical electric and magnetic fields have dimensions $[\boldsymbol{E}] = [B] = EL^{-3}$, which agrees with

$$EL^{-3} = [B^2] = [n]^2\,[\phi_0]^2 = L^{-4}\cdot\frac{[hc]^2}{[e^2]} = \frac{E^2L^{-2}}{EL} = EL^{-3} \quad . \tag{3.366}$$

[88] Note that $J^\mu = -ej^\mu$ is the *electrical current* of the Chern-Simons bosons, while $J^\mu_{\rm v}$ is the *number current* of the vortices.

3.5.7 Superfluid response and CSGL theory

The action S_{mat} in Eqn. 3.338 corresponds to the (first quantized) Hamiltonian

$$H(\delta\mathsf{a}) = \frac{1}{2m^*}\sum_{i=1}^{N}\left(\boldsymbol{p}_i + \frac{e}{c}\,\delta\vec{\mathsf{a}}_i\right)^2 - \frac{e}{c}\sum_{i=1}^{N}\delta\mathsf{a}_i^0 + \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j) \quad . \tag{3.367}$$

Here we are invoking the shifted CS gauge potential, $\delta\mathsf{a}_\mu = \delta a_\mu + \delta A_\mu$ and $A_\mu = \bar{A}_\mu + \delta A_\mu$, as in §**??**. The CS action in terms of $\delta\mathsf{a}$ and δA is given by

$$S_{\mathsf{CS}} = \frac{\pi e}{2\theta c\phi_0}\int d^3x\,\epsilon^{\mu\nu\lambda}\left\{(\delta\mathsf{a}_\mu - \delta A_\mu)\,\partial_\nu\,(\delta\mathsf{a}_\lambda - \delta A_\lambda) - 2\,\delta\mathsf{a}_\mu\,\partial_\nu\,\bar{A}_\lambda\right\} \quad . \tag{3.368}$$

Thus, we have

$$H(\delta\mathsf{a}) = H(0) - \frac{e}{c}\int d^2x\; j^{\mathsf{p}}_\mu(\boldsymbol{x})\,\delta\mathsf{a}^\mu(\boldsymbol{x}) + \frac{e^2}{2m^*c^2}\int d^2x\; n(\boldsymbol{x})\left[\delta\vec{\mathsf{a}}(\boldsymbol{x})\right]^2 \quad , \tag{3.369}$$

where j^{p}_μ is the paramagnetic current. A review of the linear response formalism is given in §3.8 below.

The effective action, once matter fields are integrated away, is given by[89]

$$\begin{aligned} S_{\mathsf{eff}}[\delta a,\delta A] &= \frac{1}{8\pi}\int d^3x\int d^3x'\;\delta\mathsf{a}^\mu(\boldsymbol{x},t)\,K_{\mu\nu}(\boldsymbol{x}-\boldsymbol{x}',t-t')\,\delta\mathsf{a}^\nu(\boldsymbol{x}',t') + S_{\mathsf{CS}}[\delta a,\delta A] \\ &= \frac{1}{8\pi}\int\frac{d^3q}{(2\pi)^3}\begin{pmatrix}\delta\mathsf{a}^\mu(-q) & \delta A^\mu(-q)\end{pmatrix}\begin{pmatrix}\hat{K}_{\mu\nu}(q)+\hat{L}_{\mu\nu}(q) & -\hat{L}_{\mu\nu}(q)\\ -\hat{L}_{\mu\nu}(q) & \hat{L}_{\mu\nu}(q)\end{pmatrix}\begin{pmatrix}\delta\mathsf{a}^\nu(q)\\ \delta A^\nu(q)\end{pmatrix} \end{aligned} \tag{3.370}$$

where $K_{\mu\nu}(\boldsymbol{x}-\boldsymbol{x}',t-t')$ is the electromagnetic response tensor derived in §3.8, and where we retain only finite wavevector components of the gauge fields. Here it proves useful to redefine $\delta\mathsf{a}^0 \to \delta\mathsf{a}^0/c$ and $\delta A^0 \to \delta A^0/c$ so that all the components of $\delta\mathsf{a}^\mu$ and δA^μ have the same dimensions. Similarly, we take $x^\mu = (ct,\boldsymbol{x})$ and $q^\mu = (c^{-1}\omega,\boldsymbol{q})$. Then

$$\hat{K}_{\mu\nu}(q) = \begin{pmatrix}(c^2\boldsymbol{q}^2/\omega^2)\,\hat{K}_\parallel(q) & (cq_j/\omega)\,\hat{K}_\parallel(q)\\ (cq_i/\omega)\,\hat{K}_\parallel(q) & \hat{q}_i\,\hat{q}_j\,\hat{K}_\parallel(q) + (\delta_{ij}-\hat{q}_i\,\hat{q}_j)\,\hat{K}_\perp(q)\end{pmatrix} \tag{3.371}$$

and

$$\hat{L}_{\mu\nu}(q) = \frac{2\pi i\alpha}{\theta c}\begin{pmatrix}0 & -cq_2 & +cq_1\\ +cq_2 & 0 & +\omega\\ -cq_1 & -\omega & 0\end{pmatrix} \quad , \tag{3.372}$$

where $\alpha = e^2/\hbar c$ is the fine structure constant. Note that both matrices are Hermitian.

[89] We only include fields at nonzero wavelength and/or frequency.

Now we integrate out $\delta \mathsf{a}^\mu$, resulting in the new reduced effective action

$$S_{\rm red}[\delta A^\mu] = \frac{1}{8\pi}\int \frac{d^3q}{(2\pi)^3}\,\delta A^\mu(-q)\left(\hat L(q) - \hat L(q)\big[\hat K(q)+\hat L(q)\big]^{-1}\hat L(q)\right)_{\mu\nu}\delta A^\nu(q) \quad . \tag{3.373}$$

To inver the matrices $\hat K_{\mu\nu}$ and $\hat L_{\mu\nu}$, it is convenient to work solely with components of q with lowered indices, since these matrices are expressed above in those variables. We define the orthonormal triad,

$$\psi_{0,\mu} = \frac{1}{\sqrt{\omega^2+c^2\boldsymbol{q}^2}}\begin{pmatrix}\omega\\ -cq_1\\ -cq_2\end{pmatrix} \quad , \quad \psi_{1,\mu} = \frac{1}{|\boldsymbol{q}|}\begin{pmatrix}0\\ -q_2\\ +q_1\end{pmatrix} \quad , \quad \psi_{2,\mu} = \frac{1}{|\boldsymbol{q}|\sqrt{\omega^2+c^2\boldsymbol{q}^2}}\begin{pmatrix}c\boldsymbol{q}^2\\ \omega q_1\\ \omega q_2\end{pmatrix} \quad , \tag{3.374}$$

which satisfy the orthogonality relations $\sum_\mu \psi_{a,\mu}(q)\,\psi_{b,\mu}(q) = \delta_{ab}$ and the completeness relations $\sum_a \psi_{a,\mu}(q)\,\psi_{a,\nu}(q) = \delta_{\mu\nu}$ for all q. Suppressing q, one readily obtains

$$\hat K\,|\,0\,\rangle = 0 \quad , \quad \hat K\,|\,1\,\rangle = \hat K_\perp\,|\,1\,\rangle \quad , \quad \hat K\,|\,2\,\rangle = \widetilde K_\parallel\,|\,2\,\rangle \quad , \tag{3.375}$$

where

$$\widetilde K_\parallel(\boldsymbol{q},\omega) \equiv \left(\frac{c^2\boldsymbol{q}^2}{\omega^2}+1\right)\hat K_\parallel(\boldsymbol{q},\omega) = -\left(1+\frac{\omega^2}{c^2\boldsymbol{q}^2}\right)4\pi e^2\,\hat\chi(\boldsymbol{q},\omega) \quad , \tag{3.376}$$

where $\hat\chi(\boldsymbol{q},\omega)$ is the scalar susceptibility of the corresponding neutral superfluid (see §3.8.3); note that $\hat\chi(\boldsymbol{q}\to 0,0) = n^2\kappa_T$ is *finite*. We also have

$$\hat L\,|\,0\,\rangle = 0 \quad , \quad \hat L\,|\,1\,\rangle = i\beta\,|\,2\,\rangle \quad , \quad \hat L\,|\,2\,\rangle = -i\beta\,|\,1\,\rangle \quad , \tag{3.377}$$

where

$$\beta(\boldsymbol{q},\omega) = \frac{2\pi\alpha}{\theta}\sqrt{\frac{\omega^2}{c^2}+\boldsymbol{q}^2} \quad . \tag{3.378}$$

Note that $|\,0\,\rangle$ is annihilated by both $\hat K$ and $\hat L$ – this is a consequence of gauge invariance. We may now write

$$\begin{aligned}\hat K &= \hat K_\perp\,|\,1\,\rangle\langle\,1\,| + \widetilde K_\parallel\,|\,2\,\rangle\langle\,2\,|\\ \hat L &= i\beta\left(|\,2\,\rangle\langle\,1\,| - |\,1\,\rangle\langle\,2\,|\right) \quad .\end{aligned} \tag{3.379}$$

Thus, in the truncated $|a\rangle$ basis ($a = 1, 2$), we have

$$\hat K + \hat L = \begin{pmatrix}\hat K_\perp & -i\beta\\ +i\beta & \widetilde K_\parallel\end{pmatrix} \quad . \tag{3.380}$$

We may now construct the pseudo-inverse

$$(\hat K+\hat L)^{-1} = \frac{1}{\hat K_\perp\widetilde K_\parallel - \beta^2}\begin{pmatrix}\widetilde K_\parallel & i\beta\\ -i\beta & \hat K_\perp\end{pmatrix} \quad , \tag{3.381}$$

and we find

$$\hat{Q} \equiv \hat{L} - \hat{L}(\hat{K} + \hat{L})^{-1}\hat{L} = -\frac{\beta^2}{\hat{K}_\perp \widetilde{K}_\parallel - \beta^2}\,\hat{K} + \frac{\hat{K}_\perp \widetilde{K}_\parallel}{\hat{K}_\perp \widetilde{K}_\parallel - \beta^2}\,\hat{L} \quad . \tag{3.382}$$

In the low frequency, long wavelength limit, $\hat{K}_\perp(\boldsymbol{q} \to 0, 0)$ and $\widetilde{K}_\parallel(\boldsymbol{q} \to 0, 0)$ are both constant and dominate over $\beta^2 \propto c^{-2}\omega^2 + \boldsymbol{q}^2$. Thus $\hat{Q} \to \hat{L}$ and we obtain the long wavelength action

$$S_{\text{red}}[\delta A^\mu] = \frac{e^2}{4\theta\hbar c}\int d^3x\, \epsilon^{\mu\nu\lambda}\,\delta A_\mu\,\partial_\nu\,\delta A_\lambda \quad , \tag{3.383}$$

exactly as in Eqn. 3.361.

Now you may ask: where are the vortices? Our description of superfluid response doesn't include them! To account for vortices, consider the vortex 3-current,

$$J^\mu_{\text{v}} = \sum_{i=1}^{N_{\text{v}}} q_i \begin{Bmatrix} 1 \\ \dot{\boldsymbol{X}}_i(t) \end{Bmatrix} \delta\big(\boldsymbol{x} - \boldsymbol{X}_i(t)\big) \quad . \tag{3.384}$$

Conservation of vorticity means $\partial_\mu J^\mu_{\text{v}} = 0$, which licenses us to define a gauge field $\mathcal{V}^\mu$ whose curl is the vortex current, *viz.*

$$J^\mu_{\text{v}} = \epsilon^{\mu\nu\lambda}\,\partial_\nu \mathcal{V}_\lambda \quad . \tag{3.385}$$

We now add a term to the Lagrangian for the superfluid particles,

$$\Delta L = 2\pi\hbar \int d^2x\, j^{\text{p}}_\mu\, \mathcal{V}^\mu \quad , \tag{3.386}$$

which in the action provides a Berry phase of $2\pi q_i$ for each time a bosonic particle of the superfluid executes a closed path encircling the i^{th} vortex[90]. In the Hamiltonian description of the superfluid, this amounts to the replacement

$$\delta\mathsf{a}^\mu \longrightarrow \delta\mathsf{a}^\mu - \phi_0 \mathcal{V}^\mu \quad . \tag{3.387}$$

With this refinement, *all* the terms in Eqn. 3.361 are recovered[91]. Ta da!

Why do we need to mess with this tedious response function formalism? Because the interaction potential $v(\boldsymbol{r} - \boldsymbol{r}')$ may be very strong at short distances. The correlations of the underlying superfluid may not be adequately described by a simple Gross-Pitaevskii $|\Psi|^4$ interaction.

[90] See D. P. Arovas and J. A. Freire, *Phys. Rev. B* **55**, 1068 (1997).

[91] See if you can trace the appearance of the first term on the RHS of Eqn. 3.361, which accounts for the Berry phase of each vortex as it winds around the background condensate.

3.5.8 Kohn mode and collective excitations

If we hold the field fixed at $B = qn\phi_0$ and set $\delta A^\mu = 0$, then the action of Eqn. 3.370 is given by

$$S_{\rm eff}[\delta a] = \frac{1}{8\pi}\int \frac{d^3q}{(2\pi)^3}\,\delta \mathsf{a}^\mu(-q)\left[\hat{K}_{\mu\nu}(q) + \hat{L}_{\mu\nu}(q)\right]\delta \mathsf{a}^\nu(q) \quad . \tag{3.388}$$

The frequencies of the elementary excitations are given by solving the equation

$$\begin{aligned} 0 &= \det\left[\hat{K} + \hat{L}\right] = \hat{K}_\perp \widetilde{K}_\parallel - \beta^2 \\ &= -\hat{K}_\perp(\boldsymbol{q},\omega)\left(1 + \frac{\omega^2}{c^2\boldsymbol{q}^2}\right)4\pi e^2\hat{\chi}(\boldsymbol{q},\omega) - \left(\frac{2\pi\alpha}{\theta}\right)^2\left(\boldsymbol{q}^2 + \frac{\omega^2}{c^2}\right) \quad , \end{aligned} \tag{3.389}$$

which says

$$4\pi e^2\,\hat{K}_\perp(\boldsymbol{q},\omega)\,\hat{\chi}(\boldsymbol{q},\omega) + \left(\frac{2\pi\alpha}{\theta}\right)^2\boldsymbol{q}^2 = 0 \quad . \tag{3.390}$$

At $T = 0$, there is no normal component to the superfluid to produce a transverse response, and we have

$$\lim_{\boldsymbol{q}\to 0}\hat{K}_\perp(\boldsymbol{q},0) = \frac{4\pi e^2 n}{m^* c^2} \quad . \tag{3.391}$$

For the density response function, we will use the SMA formula,

$$\hat{\chi}_{\rm SMA}(\boldsymbol{q},\omega) = \frac{n\boldsymbol{q}^2/m}{\omega^2(\boldsymbol{q}) - (\omega + i\epsilon)^2} \quad . \tag{3.392}$$

If $\omega(\boldsymbol{q}) = c|\boldsymbol{q}|$, we obtain the dispersion relation

$$\omega^2 = c^2\boldsymbol{q}^2 + \left(\frac{4\pi n e^2}{m^*}\right)^2\left(\frac{\theta}{2\pi\alpha}\right)^2 = \omega_{\rm c}^2 + c^2\boldsymbol{q}^2 \quad , \tag{3.393}$$

where $\omega_{\rm c} = eB/m^*c$ is the cyclotron frequency. We have found the Kohn mode.

However, the presumption of a long-wavelength dispersion $\omega(\boldsymbol{q}) = c|\boldsymbol{q}|$ for the phonons of the superfluid is not correct here, due to the long range of the interaction potential $v(\boldsymbol{r})$. As we derive in §3.7 below, the long-wavelencth phonon dispersion is rather given by

$$\omega(\boldsymbol{q}) = \sqrt{\frac{n_0\,\hat{v}(\boldsymbol{q})}{m^*}}\,|\boldsymbol{q}| \quad , \tag{3.394}$$

which for $\hat{v}(\boldsymbol{q}) = 2\pi e^2/\epsilon|\boldsymbol{q}|$ behaves as $|\boldsymbol{q}|^{1/2}$. This, you may recall, is the form for the L-phonon in the two-dimensional Wigner crystal with $1/r$ Coulomb interactions. Thus, we should expect

$$\omega^2 = \omega_{\rm c}^2 + \frac{2\pi n_0 e^2}{\epsilon m^*}\,|\boldsymbol{q}| \tag{3.395}$$

for the Kohn mode at long wavelengths.

In the original Zhang-Hansson-Kivelson paper on the CSGL theory, this inter-LL Kohn mode was misidentified as an intra-LL collective mode. This interpretation was subsequently revisited by Lee and Zhang[92], who provided a revised understanding of the intra-LL collective mode in the CSGL theory in terms of vortex-antivortex pairs and quadrupoles. If we add the long-ranged instantaneous Coulomb interaction between vortices to the theory of Eqn. 3.361, *i.e.* a term

$$\Delta S_{\rm CS} = -\frac{\pi^2}{2\theta^2}\int dt \int d^2x \int d^2x'\; J^0_{\rm v}(\boldsymbol{x},t)\, v(\boldsymbol{x}-\boldsymbol{x}')\, J^0_{\rm v}(\boldsymbol{x}',t) \quad , \tag{3.396}$$

where $v(r) = e^2/\epsilon r$, then with $\delta A^\mu = 0$ we obtain a theory of vortices and antivortices *confined to the lowest Landau level*, since there is no vortex mass term to set a scale for a vortex cyclotron energy. Interpreting the first term as measuring the enclosed area swept out by each vortex in units of l^2, where l is a vortex "magnetic length", we have $l = (2\pi n_0)^{-1/2} = \nu^{-1/2}\ell$, since the number density is given by $n_0 = \nu/2\pi\ell^2$. The remaining terms are the Coulomb interaction from $\Delta S_{\rm CS}$, and the topological term proportional to the vortex linking numbers which endows the vortices with fractional exchange statistics. As the linking number is always an integer, this can only change discontinuously due to the crossing of vortex world lines, and cannot affect the vortex equations of motion. Thus, the Lagrangian is[93]

$$S_{\rm vor} = \frac{\nu}{2\ell^2}\sum_{i=1}^{N_{\rm V}} q_i\,\epsilon_{ab}\,X_i^a\,\dot X_i^b - \frac{\nu^2 e^2}{\epsilon}\sum_{i<j}^{N_{\rm V}}\frac{q_i\,q_j}{|\boldsymbol{X}_i-\boldsymbol{X}_j|} - \sum_{i=1}^{N_{\rm V}} q_i\,U(\boldsymbol{X}_i) - \tilde\varepsilon^{\rm QE}N_{\rm QE} - \tilde\varepsilon^{\rm QH}N_{\rm QH} \quad , \tag{3.397}$$

where we have also included a one-body potential $U(\boldsymbol{r})$ for the vortices which reflects the random potential coupling to the density $n = n_0 + \delta n$, recognizing that each vortex produces a local surplus or deficit of physical electrons. Recall from §3.3.7 that a quasielectron-quasihole pair is an exciton whose wavevector $\boldsymbol{k}$ is related to the qe-qh separation $\boldsymbol{r}$ according to $\boldsymbol{k} = \nu\hat{\boldsymbol{z}}\times\boldsymbol{r}/\ell^2$. The energy of a single exciton is $\Delta_{\rm EX} = \tilde\varepsilon^{\rm QE} + \tilde\varepsilon^{\rm QH} + v(\boldsymbol{r})$ where for Coulomb interactions $v(\boldsymbol{r}) = (\nu e)^2/\epsilon r$. To create an excitation at zero wavevector, one can make a quadrupole with zero net dipole moment. This suggests that the $\boldsymbol{k} = 0$ magnetophonon is a quadrupole with energy $\Delta(\boldsymbol{0})$ should be on the order of $2\tilde\varepsilon^{\rm QE} + 2\tilde\varepsilon^{\rm QH}$, whereas the magnetoroton, which is a finite $\boldsymbol{k}$ excitation, is a dipole with energy on the order of $\tilde\varepsilon^{\rm QE} + \tilde\varepsilon^{\rm QH}$. Indeed, whereas the dipoles have a definite energy-momentum relationship, $\boldsymbol{k} = 0$ quadrupoles are available in a continuum of states. Consider a configuration with two quasielectrons at positions $\pm(x,y)$ and two quasiholes at positions $\pm(x,-y)$. For every choice of (x,y) the net dipole moment is zero, and the energy is

$$\Delta_{\rm QUAD} = 2\tilde\varepsilon^{\rm QE} + 2\tilde\varepsilon^{\rm QH} + 2v\Big(2\sqrt{x^2+y^2}\Big) - 2v(2x) - 2v(2y) \quad . \tag{3.398}$$

From numerical calculations, the $\boldsymbol{q} = 0$ portion of the collective excitation spectrum is indeed a continuum, the bottom edge of which lies above the magnetoroton minimum (see Fig. 3.10).

[92] D.-H. Lee and S.-C. Zhang, *Phys. Rev. Lett.* **66**, 1220 (1991).

[93] For a more general interaction potential, replace $e^2/\epsilon r$ with $v(r)$.

Read's version of the CSGL theory has the virtue of describing only LLL physics. His coefficient of the covariant derivative squared term is proportional to the Laplacian of a Hartree-type potential whose energy scale is set by $e^2/\epsilon\ell$. In the CSGL theory of ZHK, the coefficient is $\hbar^2/2m^*$. The electron mass m^* enters nowhere within Read's theory, which is apposite since in any LLL-projected theory we should be able to set $m^* \to 0$. However, the theory is unwieldy for other reasons and in fact does not yield a magnetoroton minimum in its collective excitation branch[94].

3.5.9 Quasi-LRO and CSGL theory

Recall that by varying the action in Eqn. 3.348 with respect to the gauge field $\delta\mathsf{A}^0$ we obtain the condition $\delta n = \pi\mathsf{b}/\theta\phi_0$, where b is the magnetic field strength corresponding to the shifted gauge field $\delta\vec{\mathsf{a}}$. Since the action is linear in $\delta\mathsf{a}^0$, this result is exact, and we may substitute it back into the remaining terms of the action with no approximations. In Fourier space, we have, taking $\theta = \pi q = \pi/\nu$,

$$\delta\hat{\mathsf{a}}^i(k) = \frac{\hbar c}{e}\,\frac{2\pi}{\nu \boldsymbol{k}^2}\,i\epsilon_{ij}\,k^j\,\delta\hat{n}(k) \quad , \tag{3.399}$$

where, following Zhang (1992), we work in the transverse gauge $\vec{\nabla}\cdot\delta\vec{\mathsf{a}} = 0$. We substitute the above result into the rest of the Lagrangian density,

$$\begin{aligned}\mathcal{L} = -\hbar\,\delta n\,\partial_t\phi - \frac{\hbar^2}{2m}\,(n_0+\delta n)\left((\boldsymbol{\nabla}\phi)^2 + \frac{2e}{\hbar c}\,\boldsymbol{\nabla}\phi\cdot\delta\vec{\mathsf{a}} + \frac{e^2}{\hbar^2c^2}\left(\delta\vec{\mathsf{a}}\right)^2\right) \\ - \frac{\hbar^2}{8m(n_0+\delta n)}\,(\boldsymbol{\nabla}\delta n)^2 - \frac{\nu e^2}{4\pi\hbar c^2}\,\epsilon_{ij}\,\delta\mathsf{a}^i\,\partial_t\,\delta\mathsf{a}^j - \tfrac{1}{2}\,\delta n\,v\,\delta n \quad ,\end{aligned} \tag{3.400}$$

where $\Psi = \sqrt{n}\,e^{i\phi}$ and where the last term is shorthand for what is in the action written as a double integral over $\boldsymbol{x}$ and $\boldsymbol{x}'$, with potential $v(\boldsymbol{x}-\boldsymbol{x}')$, as previously. *We assume that no vortices are present* and thus that $\phi = \phi_{\mathsf{sw}}$. In the transverse gauge, this means that $\boldsymbol{\nabla}\phi\cdot\delta\vec{\mathsf{a}} = 0$. We also neglect terms cubic in the density. The result of the substitution is then

$$\hat{\mathcal{L}} = i\hbar\omega\,\delta\hat{n}(-k)\,\hat{\phi}(k) - \frac{\nu}{4\pi}\,\hbar\omega_{\mathrm{c}}\,\boldsymbol{k}^2\,|\hat{\phi}(k)|^2 - \frac{1}{2}\left(\hat{v}(\boldsymbol{k}) + \frac{\hbar^2\boldsymbol{k}^2}{4mn_0} + \frac{2\pi}{\nu}\,\frac{\hbar\omega_{\mathrm{c}}}{\boldsymbol{k}^2}\right)|\delta\hat{n}(k)|^2 \quad . \tag{3.401}$$

We now integrate out the density fluctuations, using their equation of motion. Varying with respect to $\delta\hat{n}(-k)$ yields

$$\delta\hat{n}(k) = \frac{i\hbar\omega\,\hat{\phi}(k)}{\hat{v}(\boldsymbol{k}) + \frac{\hbar^2\boldsymbol{k}^2}{4mn_0} + \frac{2\pi}{\nu}\,\frac{\hbar\omega_{\mathrm{c}}}{\boldsymbol{k}^2}} \quad , \tag{3.402}$$

which results in the effective Fourier space Lagrangian density for the ϕ field,

$$\hat{\mathcal{L}}(\phi) = \frac{\nu}{4\pi}\,\hbar\omega_{\mathrm{c}}\,\boldsymbol{k}^2\left(\frac{\omega^2 - \Omega_{\mathrm{K}}^2(\boldsymbol{k})}{\Omega_{\mathrm{K}}^2(\boldsymbol{k})}\right)|\hat{\phi}(k)|^2 \quad , \tag{3.403}$$

[94] Read obtains, correctly, a magnetophonon gap on the order of $e^2/\epsilon\ell$, but upwardly dispersing as a function of wavevector.

where

$$\Omega_{\rm K}(\boldsymbol{k}) = \left[\omega_{\rm c}^2 + \frac{n_0}{m}\,\hat{v}(\boldsymbol{k})\,\boldsymbol{k}^2 + \left(\frac{\hbar\boldsymbol{k}^2}{2m}\right)^2\right]^{1/2} \tag{3.404}$$

is the frequency of the Kohn mode derived in §3.5.8 above[95].

We may now compute the expectation of the phase fluctuations at $T = 0$:

$$\left\langle|\hat{\phi}(\boldsymbol{k})|^2\right\rangle = \hbar\int\limits_0^\infty \frac{d\omega}{2\pi i}\,\frac{2\pi}{\nu}\,\frac{1}{\hbar\omega_{\rm c}\,\boldsymbol{k}^2}\,\frac{\Omega_{\rm K}^2(\boldsymbol{k})}{(\omega+i\epsilon)^2-\Omega_{\rm K}^2(\boldsymbol{k})} = \frac{\pi}{\nu}\,\frac{\Omega_{\rm K}(\boldsymbol{k})}{\omega_{\rm c}\,\boldsymbol{k}^2} \approx \frac{\pi}{\nu\,\boldsymbol{k}^2} \quad . \tag{3.405}$$

Thus,

$$\begin{aligned}\left\langle\Psi^*(\boldsymbol{r})\,\Psi(\boldsymbol{r}')\right\rangle &\simeq n_0\left\langle e^{i\phi(\boldsymbol{r}')}\,e^{-i\phi(\boldsymbol{r})}\right\rangle \\ &= n_0\,\exp\left[-\tfrac{1}{2}\left\langle\left[\phi(\boldsymbol{r})-\phi(\boldsymbol{r}')\right]^2\right\rangle\right] \quad ,\end{aligned} \tag{3.406}$$

and with

$$\left\langle\left[\phi(\boldsymbol{r})-\phi(\boldsymbol{0})\right]^2\right\rangle = \frac{2\pi}{\nu}\int\frac{d^2k}{(2\pi)^2}\,\frac{1-e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{\boldsymbol{k}^2} = \nu^{-1}\ln r \quad , \tag{3.407}$$

we recover the algebraic Girvin-MacDonald order[96],

$$\left\langle\Psi^*(\boldsymbol{r})\,\Psi(\boldsymbol{r}')\right\rangle \propto n_0\,|\boldsymbol{r}-\boldsymbol{r}'|^{-1/2\nu} \quad . \tag{3.408}$$

(Compare with Eqn. 3.325.)

3.6 Global Phase Diagram of the Quantum Hall Effect

Finally, we discuss the issue of phase transitions between different quantum Hall phases, following the "global phase diagram" picture of Kivelson, Lee, and Zhang (KLZ)[97] As we saw in chapter 2 (§2.1,7), for weak disorder, the extended single-particle states at the center of each disorder-broadened Landau level are separated in energy by a mobility gap in which all states are localized. This provided a quantum percolation picture of the IQH transition, which could be investigated via the Chalker-Coddington network or disordered Hoftstadter models, and when disorder is increased, the extended states "float up" in energy. Still, within this picture all direct IQH transitions involve $\Delta\sigma_{xy} = \pm e^2/h$. Despite some problems with noninteracting

[95] We have included the term proportional to $|\boldsymbol{k}|^4$ in $\Omega_{\rm K}^2(\boldsymbol{k})$ here for completeness. This arises from the $(\boldsymbol{\nabla}\delta n)^2$ term in the original Lagrangian.

[96] The integral in Eqn. 3.407 diverges in the ultraviolet and a cutoff at $k \approx \ell^{-1}$ must be imposed.

[97] S. Kivelson, D.-H. Lee, and S.-C. Zhang, *Phys. Rev. B* **46**, 2223 (1992).

models of the IQHE[98], if we ignore LL mixing then the transition between $\nu = n$ and $\nu = n+1$ plateaus occurs for $\nu = n + \frac{1}{2}$, *i.e.* in the center of each disorder-broadened LL. The transition is marked by a crossing of $\sigma_{xy}(B, T, L)$ curves at the value $\sigma_{xy} = (n + \frac{1}{2})\, e^2/\hbar$. Recall the relations, valid in isotropic systems,

$$\rho = \begin{pmatrix} \rho_{xx} & \rho_{xy} \\ -\rho_{xy} & \rho_{xx} \end{pmatrix} = \frac{1}{\sigma_{xx}^2 + \sigma_{xy}^2} \begin{pmatrix} \sigma_{xx} & -\sigma_{xy} \\ \sigma_{xy} & \sigma_{xx} \end{pmatrix} = \sigma^{-1} \tag{3.409}$$

and

$$\sigma = \begin{pmatrix} \sigma_{xx} & \sigma_{xy} \\ -\sigma_{xy} & \sigma_{xx} \end{pmatrix} = \frac{1}{\rho_{xx}^2 + \rho_{xy}^2} \begin{pmatrix} \rho_{xx} & -\rho_{xy} \\ \rho_{xy} & \rho_{xx} \end{pmatrix} = \rho^{-1} \quad . \tag{3.410}$$

Thus the $n \to n+1$ IQH transition lying at $\sigma_{xy} = (n + \frac{1}{2})\, e^2/\hbar$ entails a relation between the longitudinal and transverse components of the resistivity, $\rho_{xx} \equiv r\, h/e^2$ and $\rho_{yx} \equiv s\, h/e^2$:

$$n \text{ to } n+1 \quad : \quad r^2 + \left(s - \frac{1}{2n+1} \right)^2 = \frac{1}{(2n+1)^2} \quad . \tag{3.411}$$

In the quadrant $(s > 0, r > 0)$ of the (s, r) plane, for each n the above equation describes a half-circle, centered at $(s, r) = \left(\frac{1}{2n+1}\, ,\, 0\right)$, of radius $\frac{1}{2n+1}$. The maximum value for each n occurs at the center, and is given by $r^* = \frac{1}{2n+1}$.

The KLZ picture is based on a "law of corresponding states" which posits that the physics of a QH state at filling fraction ν is related to that at other fillings related by the operations of LL addition $(\nu \to \nu + 1)$, particle-hole transformation within the LLL $(\nu \to 1 - \nu)$, and flux addition $(\nu^{-1} \to \nu^{-1} + 2)$. We discussed these operations toward the end of §3.4.4, and one can define explicit mappings at the level of wavefunctions for each of them. KLZ provide a nonrigorous but well-motivated argument for this based on the CSGL theory. It is important to note that their procedure accommodates disorder as well.

It is convenient to define the dimensionless components of the conductivity tensor u and v according to $\sigma_{xx} \equiv u\, e^2/h$ and $\sigma_{xy} \equiv v\, e^2/h$. Thus,

$$r = \frac{u}{u^2 + v^2} \quad , \quad s = \frac{v}{u^2 + v^2} \quad , \quad u = \frac{r}{r^2 + s^2} \quad , \quad v = \frac{s}{r^2 + s^2} \quad . \tag{3.412}$$

Suppose a phase boundary between QH states at fillings ν and ν' is expressed as a relation between r and s as $F(\nu\, ,\, \nu' \,|\, r\, ,\, s) = 0$, with $\lambda F(\nu\, ,\, \nu' \,|\, r\, ,\, s) \cong F(\nu\, ,\, \nu' \,|\, r\, ,\, s)$, *i.e.* multiplication by a constant does not change the condition $F(\nu\, ,\, \nu' \,|\, r\, ,\, s) = 0$. Eqn. 3.411 may be written as

$$F\big(n\, ,\, n+1 \,|\, r\, ,\, s\big) = r^2 + s^2 - \frac{2s}{2n+1} \quad , \tag{3.413}$$

From this expression we may derive the phase boundaries for all other QH transitions within the KLZ scheme via a combination of the following operations:

[98] Recall that such noninteracting descriptions apparently don't properly recover the experimentally observed value for the correlation length exponent $\nu \approx 2.35$ and instead give $\nu \approx 2.58$; the difference is now large enough to rule out the noninteracting theory, which also fails to give $z = 1$ for the dynamic critical exponent, which is what experimental scaling analysis supports.

(i) *Landau level addition* : Under LL addition, one has $\nu = \nu_0 + 1$, $u = u_0$, and $v = v_0 + 1$. Suppose we do this n times, so $\nu = \nu_0 + n$, $u = u_0$ and $v = v_0 + n$. Then with

$$u_0 = u = \frac{r}{r^2+s^2} \quad , \quad v_0 = v - n = \frac{s - n(r^2+s^2)}{r^2+s^2} \quad , \tag{3.414}$$

we have

$$\begin{aligned} r_0^{\text{LLA}}(r,s) &= \frac{r(r^2+s^2)}{r^2 + \big[s - n(r^2+s^2)\big]^2} \\ s_0^{\text{LLA}}(r,s) &= \frac{\big[s - n(r^2+s^2)\big](r^2+s^2)}{r^2 + \big[s - n(r^2+s^2)\big]^2} \end{aligned} \tag{3.415}$$

and we may write

$$F\big(\nu + n\,,\,\nu' + n \,|\, r\,,\,s\big) = F\big(\nu\,,\,\nu' \,|\, r_0^{\text{LLA}}(r,s)\,,\,s_0^{\text{LLA}}(r,s)\big) \quad . \tag{3.416}$$

If we start with the $0 \to 1$ transition, where we may take $F_0(r,s) = r^2 + s^2 - 2s$, then we obtain

$$F\big(n\,,\,n+1 \,|\, r\,,\,s\big) = \frac{r^2+s^2}{r^2 + \big[s - n(r^2+s^2)\big]^2}\,\big[(2n+1)(r^2+s^2) - 2s\big] \quad , \tag{3.417}$$

which is congruent to the form in Eqn. 3.413.

(ii) *Particle-hole conjugation* : Under the PHC operation, assuming $\nu < 1$, one has $\nu = 1 - \nu_0$, $u = u_0$, and $v = 1 - v_0$. Suppose we do this n times, so $\nu = \nu_0 + n$, $u = u_0$, and $v = v_0 + n$. Thus

$$\begin{aligned} r_0^{\text{PHC}}(r,s) &= \frac{r(r^2+s^2)}{r^2 + \big[s - r^2 - s^2\big)\big]^2} \\ s_0^{\text{PHC}}(r,s) &= \frac{\big[r^2 + s^2 - s\big](r^2+s^2)}{r^2 + \big[s - r^2 - s^2\big]^2} \end{aligned} \tag{3.418}$$

and we may write

$$F\big(1-\nu'\,,\,1-\nu \,|\, r\,,\,s\big) = F\big(1-\nu'\,,\,1-\nu \,|\, r_0^{\text{PHC}}(r,s)\,,\,s_0^{\text{PHC}}(r,s)\big) \tag{3.419}$$

(iii) *Flux attachment* : Under the flux attachment operation, one has $\nu^{-1} = \nu_0^{-1} + 2$, $r = r_0$, and $s = s_0 + 2$. Suppose we do this p times, so $\nu = \nu_0/(2p\nu_0 + 1)$, $r = r_0$, and $s = s_0 + 2p$. Then

$$F\left(\frac{\nu}{2p\nu+1}\,,\,\frac{\nu'}{2p\nu'+1}\,\middle|\,r\,,\,s\right) = F\big(\nu\,,\,\nu' \,|\, r\,,\,s - 2p\big) \quad . \tag{3.420}$$

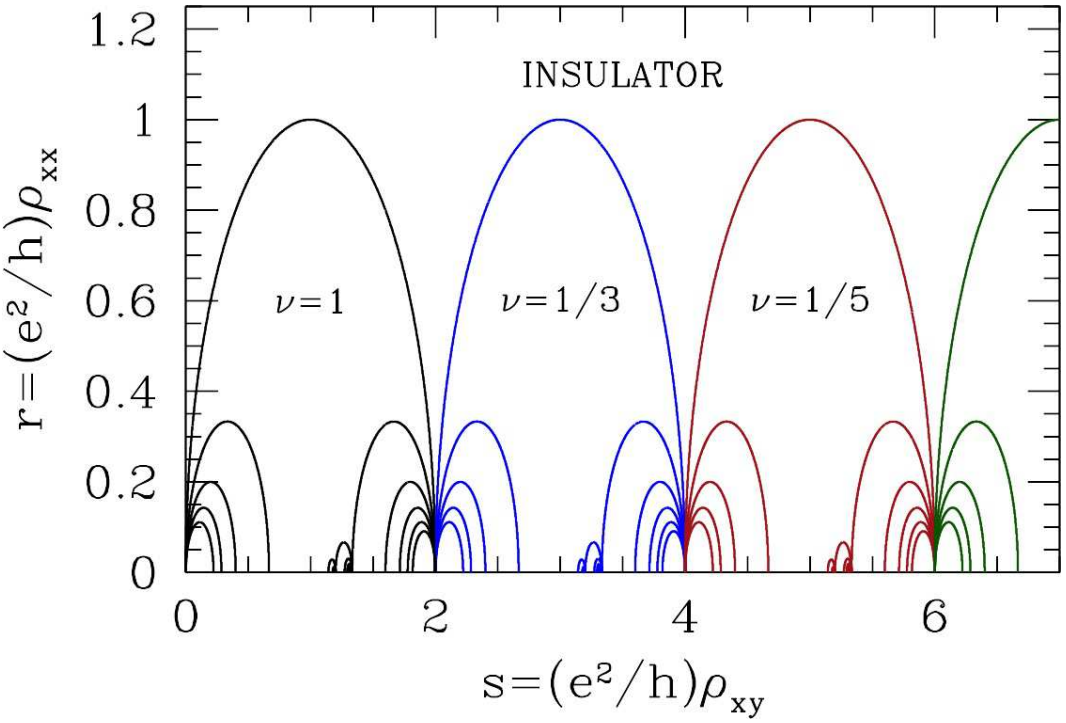


Figure 3.13: Conjectured global phase diagram of the quantum Hall effect.

Thus, we have

$$F\left(\frac{n}{2pn+1}\,,\,\frac{n+1}{2p(n+1)+1}\,\middle|\,r\,,\,s\right) = r^2 + \big(s - s_{n,p}\big)^2 - a_{n,p}^2 \tag{3.421}$$

with

$$s_{n,p} = 2p + \frac{1}{2n+1} \qquad , \qquad a_{n,p} = \frac{1}{2n+1} \quad . \tag{3.422}$$

It is a useful exercise to compute the effect of the PHC operation on the $\nu = n/(2np+1)$. After a straightforward but slightly tedious calculation, one obtains

$$F\left(\frac{(2p-1)(n+1)+1}{2p(n+1)+1}\,,\,\frac{(2p-1)n+1}{2pn+1}\,\middle|\,r,s\right) = r^2 + \big(s - \tilde{s}_{n,p}\big)^2 - \tilde{a}_{n,p}^2 \quad , \tag{3.423}$$

where

$$\tilde{s}_{n,p} = \frac{2p(2p-1)(2n+1) + 4p - 1}{(2p-1)\big[(2p-1)(2n+1)+2\big]} \tag{3.424}$$

and

$$\tilde{a}_{n,p}^2 = s_{n,p}^2 - \frac{4p\big[p(2n+1)+1\big]}{(2p-1)\big[(2p-1)(2n+1)+2\big]} \quad . \tag{3.425}$$

For the case $n = 1$ and $p = 1$, for example, we find

$$F\big(\tfrac{3}{5}\,,\,\tfrac{2}{3}\,\big|\,r,s\big) = r^2 + \big(s - \tfrac{9}{5}\big)^2 - \tfrac{1}{25} \quad . \tag{3.426}$$

A few iterations of the various operations yields the phase diagram shown in Fig. 3.13.

As KLZ stress, one should not take the details of Fig. 3.13 too seriously. Each of the three transformations results in a different effective disorder potential. Rather, it is the topology of the global phase diagram which is alleged to be robust. This tells us that there are direct IQH transitions only from $n \to n \pm 1$ and never with $\Delta n > 1$. Similarly, one may observe the transitions $\frac{1}{3} \leftrightarrow \frac{2}{5} \leftrightarrow \frac{3}{7}$, but not $\frac{1}{3} \leftrightarrow \frac{3}{7}$. Another feature of the global phase diagram is the singularities at even denominators, where there is a confluence of an infinite number of phases. In general, disorder will kill off all but a relatively small number of these phases, but rather than the insulating state extending down to the $r = 0$ axis in Fig. 3.13 at even values of s, at some point the Fermi liquid like physics of the $\nu = \frac{1}{2}$ *etc.* states sets in. We shall discuss the half-filled Landau level in the next chapter.

The insulating phase in Fig. 3.13 is identified as a *Hall insulator*, in which $\sigma_{xx} \to 0$, $\sigma_{xy} \to 0$, $\rho_{xx} \to \infty$, but $\rho_{yz} < \infty$ is a constant value roughly given by B/nec. It differs from the band insulator and Mott insulator phases, where $\rho_{yx} \to \infty$. The disordered Wigner crystal phase with finite size Imry-Ma domains could be a Hall insulator.

3.7 Appendix I: Density Correlations in a Superfluid

As a model of a vanilla superfluid, consider the Gross-Pitaevskii field theory, with Euclidean Lagrangian density

$$\mathcal{L}_{\mathsf{E}} = \hbar\,\bar{\psi}\,\partial_\tau\,\psi + \frac{\hbar^2}{2m}\,\big|\boldsymbol{\nabla}\psi\big|^2 + \tfrac{1}{2}g\big(|\psi|^2 - n_0\big) \tag{3.427}$$

in $d = 2$ space dimensions. Write $\psi = \sqrt{n}\,\exp(i\phi)$, so that

$$\mathcal{L}_{\mathsf{E}} = i\hbar\,n\,\partial_\tau\phi + \frac{\hbar^2 n}{2m}\,(\boldsymbol{\nabla}\phi)^2 + \frac{\hbar^2}{8mn}\,(\boldsymbol{\nabla}n)^2 + \tfrac{1}{2}g\,(n - n_0)^2 \quad . \tag{3.428}$$

We write $n = n_0 + \delta n$ and expand in the small quantities δn, $\boldsymbol{\nabla}\delta n$, and $\boldsymbol{\nabla}\phi$, and adding a source term j with respect to which we may differentiate. Thus,

$$\mathcal{L}_{\mathsf{E}} = i\hbar\,n_0\,\partial_\tau\phi + i\hbar\,\delta n\,\partial_\tau\phi + \frac{\hbar^2 n_0}{2m}\,(\boldsymbol{\nabla}\phi)^2 + \frac{\hbar^2}{8mn_0}\,(\boldsymbol{\nabla}\delta n)^2 + \tfrac{1}{2}g\,(\delta n)^2 + j\,\delta n \quad , \tag{3.429}$$

which is valid to quadratic order in small quantities. The first term on the RHS is important when vortices are present. Else, since it is a total derivative, in the action it integrates to zero.

We shall be interested in the case when there are no vortices, so we will drop this term. Going now to Fourier space, we have

$$\widetilde{\mathcal{L}}_{\mathsf{E}} = \Big(\hbar\omega\,\hat\phi(\boldsymbol{k},\omega) + \hat\jmath(\boldsymbol{k},\omega)\Big)\,\delta\hat n(-\boldsymbol{k},-\omega) + n_0\,\varepsilon_{\boldsymbol{k}}\,\big|\hat\phi(\boldsymbol{k},\omega)\big|^2 + \left(\frac{\varepsilon_{\boldsymbol{k}}}{4n_0} + \tfrac{1}{2}g\right)\big|\delta\hat n(\boldsymbol{k},\omega)\big|^2 \tag{3.430}$$

where $\varepsilon_{\boldsymbol{k}} = \hbar^2\boldsymbol{k}^2/2m$. Now vary the action with respect to $\delta\hat n^*(\boldsymbol{k},\omega) = \delta\hat n(-\boldsymbol{k},-\omega)$ to obtain

$$\left(\frac{\varepsilon_{\boldsymbol{k}}}{2n_0} + g\right)\delta\hat n(\boldsymbol{k},\omega) + \Big(\hbar\omega\,\hat\phi(\boldsymbol{k},\omega) + \hat\jmath(\boldsymbol{k},\omega)\Big) = 0 \quad . \tag{3.431}$$

Integrating out the density fluctuations using the above equation of motion, we obtain the Lagrangian density

$$\begin{aligned}\widetilde{\mathcal{L}}_{\mathsf{E}} = n_0\left(\frac{(\hbar\omega)^2}{\varepsilon_{\boldsymbol{k}} + 2gn_0} + \varepsilon_{\boldsymbol{k}}\right)\big|\hat\phi(\boldsymbol{k},\omega)\big|^2 - \frac{n_0\,|\hat\jmath\,(\boldsymbol{k},\omega)|^2}{\varepsilon_{\boldsymbol{k}} + 2gn_0} \\ + \frac{\hbar\omega n_0}{\varepsilon_{\boldsymbol{k}} + 2gn_0}\Big(\hat\phi^*(\boldsymbol{k},\omega)\,\hat\jmath\,(\boldsymbol{k},\omega) - \hat\phi(\boldsymbol{k},\omega)\,\hat\jmath^*(\boldsymbol{k},\omega)\Big) \quad .\end{aligned} \tag{3.432}$$

From the coefficient of the $|\hat\phi|^2$ term we can read off the phonon dispersion,

$$\omega(\boldsymbol{k}) = \frac{1}{\hbar}\sqrt{\varepsilon_{\boldsymbol{k}}\,(\varepsilon_{\boldsymbol{k}} + 2gn_0)} \quad , \tag{3.433}$$

which shows up as a pole in the $\hat\phi$ propagator at $\omega = i\omega(\boldsymbol{k})$[99]. As $k \to 0$ we obtain $\omega(\boldsymbol{k}) = u|\boldsymbol{k}|$ with a phonon velocity $u = \sqrt{gn_0/m}$. As expected, in the ultraviolet limit $k \to \infty$ we recover the ballistic dispersion $\omega(\boldsymbol{k}) = \varepsilon_{\boldsymbol{k}}/\hbar$.

We can now integrate out the phase fluctuations $\hat\phi(\boldsymbol{k},\omega)$ using the same method to obtain the Euclidean action as a function of the source $\hat\jmath(\boldsymbol{k},\omega)$:

$$S_{\mathsf{E}}[j] = -\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\int\frac{d^2k}{(2\pi)^2}\,\frac{n_0\,\varepsilon_{\boldsymbol{k}}}{(\hbar\omega)^2 + \varepsilon_{\boldsymbol{k}}(\varepsilon_{\boldsymbol{k}} + 2gn_0)}\,\big|\hat\jmath(\boldsymbol{k},\omega)\big|^2 \quad . \tag{3.434}$$

In essence we have just done two Gaussian functional integrals. As a sanity check, note that setting $m \to \infty$ and then $\hbar \to 0$ kills off all but the last two terms on the RHS of Eqn. 3.429, and accordingly in this limit the integrand becomes $|\hat\jmath|^2/2g$.

Differentiating now with respect to the source, we obtain the equal-time correlator for density fluctuations

$$\begin{aligned}s(\boldsymbol{k}) &= \frac{1}{n_0}\int d^2r\;e^{i\boldsymbol{k}\cdot\boldsymbol{r}}\,\big\langle\delta n(\boldsymbol{r},0)\,\delta n(\boldsymbol{0},0)\big\rangle \\ &= -\frac{\hbar}{n_0}\int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{\delta^2 S[j]}{\delta\hat\jmath^*(\boldsymbol{k},\omega)\,\delta\hat\jmath(\boldsymbol{k},\omega)} = \int_{-\infty}^{\infty}\frac{d\omega}{2\pi}\,\frac{2\hbar\,\varepsilon_{\boldsymbol{k}}}{(\hbar\omega)^2 + \varepsilon_{\boldsymbol{k}}(\varepsilon_{\boldsymbol{k}} + 2gn_0)} = \sqrt{\frac{\varepsilon_{\boldsymbol{k}}}{\varepsilon_{\boldsymbol{k}} + 2gn_0}} \quad .\end{aligned} \tag{3.435}$$

[99] Had we been working in real time, rather than Euclidean time, the pole would have been at $\omega = \omega(\boldsymbol{k})$.

At long wavelengths we have $s(\boldsymbol{k}) = \frac{1}{\sqrt{2}}\, k\lambda$ where $\lambda = \hbar/mu$ is the Compton wavelength (with u the speed of sound in the superfluid). In the ultraviolet limit, $s(\boldsymbol{k}) \to 1$ as is always the case[100]. Note that there is no superfluidity at any finite temperature $T > 0$ in $d = 2$, as a consequence of the Hohenberg-Mermin-Wagner theorem. Rather, for a model with $\mathsf{O}(2)$ symmetry, there is a Kosterlitz-Thouless phase transition at a critical temperature T_{KT}, below which thermally excited vortices and antivortices are bound. For $T < T_{\mathsf{KT}}$, there is a finite superfluid stiffness $\rho_{\rm s}$, but the condensate fraction n_0 is rigorously zero, in accordance with the HMW theorem[101]. For $T > T_{\mathsf{KT}}$, there is a vortex-antivortex plasma.

As we've seen, the dispersion $\omega(\boldsymbol{k})$ in the Gross-Pitaevskii model crosses over from the acoustic phonon behavior $u|\boldsymbol{k}|$ in the infrared to the ballistic $\hbar^2\boldsymbol{k}^2/2m$ in the ultraviolet. It is thus a convex function of k and shows no hint of a roton dip. The reason is that the contact interaction $v(\boldsymbol{r} - \boldsymbol{r}') = g\,\delta(\boldsymbol{r} - \boldsymbol{r}')$ is purely repulsive. If we replace it with a more general interaction potential $g(\boldsymbol{r} - \boldsymbol{r}')$, then we can accommodate a roton in our model through the behavior of its Fourier transform $\hat{g}(\boldsymbol{k})$. The GP dispersion and structure factor are then given by

$$\omega(\boldsymbol{k}) = \frac{1}{\hbar}\sqrt{\varepsilon_{\boldsymbol{k}}\,(\varepsilon_{\boldsymbol{k}} + 2\,\hat{g}(\boldsymbol{k})\,n_0)} \qquad , \qquad s(\boldsymbol{k}) = \sqrt{\frac{\varepsilon_{\boldsymbol{k}}}{\varepsilon_{\boldsymbol{k}} + 2\,\hat{g}(\boldsymbol{k})\,n_0}} \quad . \tag{3.436}$$

Thus, a phonon-roton dispersion curve is modeled if $\hat{g}(\boldsymbol{k})$ has a pronounced dip in the vicinity of $k \approx k_{\mathsf{R}}$.

3.8 Appendix II: Linear Response and Correlation Functions

We now present a litany of useful definitions and results which are applied to CSGL theory of the FQHE in §3.5.7. The Hamiltonian for a system of particles of charge $(-e)$ is given by

$$H(\mathcal{A}^\mu) = \frac{1}{2m^*}\sum_{i=1}^N \Big(\boldsymbol{p}_i + \frac{e}{c}\,\boldsymbol{\mathcal{A}}_i\Big)^2 - \frac{e}{c}\sum_{i=1}^N \mathcal{A}_i^0 + \sum_{i<j} v(\boldsymbol{r}_i - \boldsymbol{r}_j) \quad . \tag{3.437}$$

We begin with the definitions

$$\begin{aligned} j_0^{\mathsf{P}}(\boldsymbol{x}) &= n(\boldsymbol{x}) = \sum_{i=1}^N \delta(\boldsymbol{x} - \boldsymbol{x}_i) \\ \boldsymbol{j}^{\mathsf{P}}(\boldsymbol{x}) &= \frac{1}{2m^*}\sum_{i=1}^N \Big[\boldsymbol{p}_i\,\delta(\boldsymbol{x} - \boldsymbol{x}_i) + \delta(\boldsymbol{x} - \boldsymbol{x}_i)\,\boldsymbol{p}_i\Big] \quad , \end{aligned} \tag{3.438}$$

[100] Recall $s(\boldsymbol{k}) = N^{-1}\sum_{i,j}\langle e^{i\boldsymbol{k}\cdot\boldsymbol{r}_i}\,e^{-i\boldsymbol{k}\cdot\boldsymbol{r}_j}\rangle$. For $k \to \infty$, only terms with $i = j$ contribute and $s(\boldsymbol{k}) \to 1$.

[101] The superfluid stiffness $\rho_{\rm s}(T)$ vanishes for $T > T_{\mathsf{KT}}$, but rather than vanishing as a power law, as in conventional second order phase transitions, it exhibits a *universal jump* such that $\rho_{\rm s}(T_{\mathsf{KT}}^-) = 2k_{\rm B}T_{\mathsf{KT}}/\pi$.

where $j^{\rm p}_\mu$ is called the *paramagnetic number current*[102]. For charged systems in the presence of an electromagnetic field with vector potential $\mathcal{A}^\mu(\boldsymbol{x},t)$, there is also a *diamagnetic number current*,

$$j^{\rm d}_0 = 0 \quad , \quad \boldsymbol{j}^{\rm d}(\boldsymbol{x}) = \frac{e}{m^*c}\, n(\boldsymbol{x})\, \boldsymbol{\mathcal{A}}(\boldsymbol{x}) \quad , \tag{3.439}$$

where the particle charge is $(-e)$, and the *gauge-invariant current operator* is given by

$$j_\mu = -\frac{c}{e}\,\frac{\delta H}{\delta \mathcal{A}^\mu} = j^{\rm p}_\mu + j^{\rm d}_\mu \quad . \tag{3.440}$$

3.8.1 Linear response theory

Consider a quantum Hamiltonian $H(t) = H_0 - \sum_i Q_i\,\phi_i(t)$ where $\{Q_i\}$ are operators and the $\{\phi_i(t)\}$ are fields or potentials[103]. From first order perturbation theory, one derives the linear response relation

$$\big\langle\, Q(t)\,\big\rangle = \int\limits_{-\infty}^{\infty} dt'\,\chi_{ij}(t-t')\,\phi_j(t') + \mathcal{O}\big(\phi^2\big) \quad , \tag{3.441}$$

where the *response functions* $\chi_{ij}(t-t')$ are given by

$$\chi_{ij}(t-t') = \frac{i}{\hbar}\,\big\langle \big[Q_i(t), Q_j(t')\big]\big\rangle\,\Theta(t-t') \quad , \tag{3.442}$$

where $\langle\cdots\rangle$ denotes a thermal average. Thus, the Fourier transform $\hat\chi_{ij}(\omega)$ is given by

$$\begin{aligned}\hat\chi_{ij}(\omega) &= \frac{i}{\hbar}\int\limits_0^\infty dt\,\big\langle\big[Q_i(t),Q_j(0)\big]\big\rangle\,e^{i\omega t} \\ &= \frac{1}{\hbar Z}\sum_{m,n} e^{-\beta E_m}\left\{\frac{\langle\, m\,|\,Q_j\,|\,n\,\rangle\,\langle\, n\,|\,Q_i\,|\,m\,\rangle}{\omega - E_m + E_n + i\epsilon} - \frac{\langle\, m\,|\,Q_i\,|\,n\,\rangle\,\langle\, n\,|\,Q_j\,|\,m\,\rangle}{\omega + E_m - E_n + i\epsilon}\right\} \quad .\end{aligned} \tag{3.443}$$

At $T=0$, with $\omega_n \equiv E_n - E_0$, we have

$$\hat\chi_{ij}(\omega) = \frac{1}{\hbar}\sum_n\left\{\frac{\langle\, 0\,|\,Q_j\,|\,n\,\rangle\,\langle\, n\,|\,Q_i\,|\,0\,\rangle}{\omega + \omega_n + i\epsilon} - \frac{\langle\, 0\,|\,Q_i\,|\,n\,\rangle\,\langle\, n\,|\,Q_j\,|\,0\,\rangle}{\omega - \omega_n + i\epsilon}\right\} \quad . \tag{3.444}$$

[102] Note that with our $(+--)$ metric that $V^0 = V_0$ for any 3-vector V^μ, and hence $(j^{\rm p})^0 = j^{\rm p}_0$.
[103] We also assume that $\langle Q_i\rangle = 0$ when all the ϕ_i are set to zero.

3.8.2 Electromagnetic response

In the case of electromagnetic response for charge $(-e)$ objects, we define $J^{\mathsf{p}}_\mu \equiv -ej^{\mathsf{p}}_\mu$. In the presence of an electromagnetic vector potential $\mathcal{A}^\mu$ we have

$$H(\mathcal{A}) = H(0) + \frac{1}{c}\int d^2x\; J^{\mathsf{p}}_\mu\,\mathcal{A}^\mu - \frac{e}{2m^*c^2}\int d^2x\; J^{\mathsf{p}}_0\,\boldsymbol{\mathcal{A}}^2 \quad . \tag{3.445}$$

Linear response theory then says

$$\big\langle J_\mu(\boldsymbol{x},t)\big\rangle = \frac{c}{4\pi}\int d^2x'\; K_{\mu\nu}(\boldsymbol{x}-\boldsymbol{x}',t-t')\,\mathcal{A}^\nu(\boldsymbol{x}',t') \quad , \tag{3.446}$$

where the electromagnetic response tensor is

$$K_{\mu\nu}(\boldsymbol{x}-\boldsymbol{x}',t-t') = \frac{4\pi}{i\hbar c^2}\left\langle\big[J^{\mathsf{p}}_\mu(\boldsymbol{x},t),J^{\mathsf{p}}_\nu(\boldsymbol{x}',t')\big]\right\rangle\Theta(t-t') + \frac{4\pi e}{m^*c^2}\left\langle J^{\mathsf{p}}_0(\boldsymbol{x})\right\rangle g_{\mu\nu}\,(1-\delta_{\mu0})\,\delta(\boldsymbol{x}-\boldsymbol{x}')\,\delta(t-t') \quad , \tag{3.447}$$

where, recall, the metric is $g = \mathsf{diag}(+,-,-)$. Taking the Fourier transform,

$$\hat{K}_{\mu\nu}(\boldsymbol{q},\omega) = \frac{4\pi e^2}{i\hbar c^2}\int_0^\infty dt\,\left\langle\big[\hat{\jmath}^{\mathsf{p}}_\mu(\boldsymbol{q},t),\hat{\jmath}^{\mathsf{p}}_\nu(-\boldsymbol{q},0)\big]\right\rangle e^{i\omega t} + \frac{4\pi e^2}{c^2}\,\frac{n}{m^*}\,\delta_{\mu\nu}(1-\delta_{\mu0}) \quad . \tag{3.448}$$

Linear response says

$$\big\langle J_\mu(\boldsymbol{q},\omega)\big\rangle = \frac{c}{4\pi}\,\hat{K}_{\mu\nu}(\boldsymbol{q},\omega)\,\mathcal{A}^\nu(\boldsymbol{q},\omega) \quad . \tag{3.449}$$

Gauge invariance requires that the physical current J_μ is unchanged if $\mathcal{A}^\mu \to \mathcal{A}^\mu + \partial^\mu f$. Charge conservation requires $\partial^\mu J_\mu = 0$. These two conditions therefore guarantee

$$q^\mu\,\hat{K}_{\mu\nu}(\boldsymbol{q},\omega) = \hat{K}_{\mu\nu}(\boldsymbol{q},\omega)\,q^\nu = 0 \quad , \tag{3.450}$$

with $q^\mu = (\omega,\boldsymbol{q})$. In fact, these two conditions are equivalent, as a consequence of Onsager reciprocity, which guarantees

$$\begin{aligned}\mathsf{Re}\,\hat{K}_{\mu\nu}(\boldsymbol{q},\omega) &= +\mathsf{Re}\,\hat{K}_{\nu\mu}(-\boldsymbol{q},-\omega)\\ \mathsf{Im}\,\hat{K}_{\mu\nu}(\boldsymbol{q},\omega) &= -\mathsf{Im}\,\hat{K}_{\nu\mu}(-\boldsymbol{q},-\omega) \quad .\end{aligned} \tag{3.451}$$

Furthermore, spatial isotropy requires that

$$\hat{K}_{ij}(\boldsymbol{q},\omega) = \hat{K}_\parallel(\boldsymbol{q},\omega)\,\hat{q}_i\,\hat{q}_j + \hat{K}_\perp(\boldsymbol{q},\omega)\,(\delta_{ij}-\hat{q}_i\,\hat{q}_j) \quad , \tag{3.452}$$

where $\hat{\boldsymbol{q}} = \boldsymbol{q}/|\boldsymbol{q}|$. We may now invoke gauge invariance and charge conservation to establish

$$\begin{aligned}\hat{K}_{0j}(\boldsymbol{q},\omega) = \hat{K}_{j0}(\boldsymbol{q},\omega) &= \frac{q_j}{\omega}\,\hat{K}_\parallel(\boldsymbol{q},\omega)\\ \hat{K}_{00}(\boldsymbol{q},\omega) &= \frac{\boldsymbol{q}^2}{\omega^2}\,\hat{K}_\parallel(\boldsymbol{q},\omega) \quad .\end{aligned} \tag{3.453}$$

If we choose the gauge $\mathcal{A}^0 = 0$, then $\boldsymbol{\mathcal{E}} = -c^{-1}\dot{\boldsymbol{\mathcal{A}}} = i\omega c^{-1}\boldsymbol{\mathcal{A}}$ and the conductivity tensor is given by

$$\hat{\sigma}_{ij}(\boldsymbol{q},\omega) = \frac{c^2}{4\pi i\omega}\,\hat{K}_{ij}(\boldsymbol{q},\omega) \quad . \tag{3.454}$$

If $\hat{\sigma}(\boldsymbol{q},0)$ is not to diverge, when we must have $\hat{K}_{ij}(\boldsymbol{q},0) = 0$, which implies the following sum rule:

$$\frac{i}{\hbar}\int\limits_0^\infty dt\,\big\langle\big[J_i^{\mathsf{p}}(\boldsymbol{x},t), J_j^{\mathsf{p}}(0,0)\big]\big\rangle = -\frac{ne^2}{m^*}\,\delta(\boldsymbol{x})\,\delta_{ij} \quad , \tag{3.455}$$

where $n = \langle n(\boldsymbol{x})\rangle$ is presumed constant. This sum rule is violated in superconductors.

3.8.3 Neutral systems

In neutral systems, linear response theory may be applied to the particle 3-current, $j^\mu = (n, \boldsymbol{j})$, and we define the susceptibility matrix $\chi_{\mu\nu}(\boldsymbol{x},t)$ as

$$\chi_{\mu\nu}(\boldsymbol{x},t) = \frac{i}{\hbar}\,\big\langle\big[j_\mu(\boldsymbol{x},t), j_\nu(0,0)\big]\big\rangle\,\Theta(t) = \begin{pmatrix}\chi_{00} & \chi_{0j}\\ \chi_{i0} & \chi_{ij}\end{pmatrix} \quad . \tag{3.456}$$

The component $\chi_{00}(\boldsymbol{x},t) \equiv \chi(\boldsymbol{x},t)$ describes the density response to a potential $U(\boldsymbol{x},t)$, with perturbing Hamiltonian $H_1 = -\int d^2x\, n(\boldsymbol{x})\,U(\boldsymbol{x},t)$. The Fourier transform is written as

$$\hat{\chi}_{\mu\nu}(\boldsymbol{q},\omega) = \int\limits_{-\infty}^{\infty} dt \int d^2x\,\chi_{\mu\nu}(\boldsymbol{x},t)\,e^{-i\boldsymbol{q}\cdot\boldsymbol{x}}\,e^{i\omega t} \quad . \tag{3.457}$$

Note that

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T = \frac{1}{n^2}\left(\frac{\partial n}{\partial \mu}\right)_T = n^{-2}\,\hat{\chi}(\boldsymbol{q}\to 0, 0) \quad . \tag{3.458}$$

Spatial isotropy says that we may write the spatial tensor

$$\hat{\chi}_{ij}(\boldsymbol{q},\omega) \equiv \hat{\chi}_\parallel(\boldsymbol{q},\omega)\,\hat{q}_i\,\hat{q}_j + \hat{\chi}_\perp(\boldsymbol{q},\omega)\,(\delta_{ij} - \hat{q}_i\,\hat{q}_j) \quad . \tag{3.459}$$

Continuity $\partial_t\, n + \boldsymbol{\nabla}\cdot\boldsymbol{j} = 0$ then guarantees

$$\begin{aligned}\hat{\chi}_{00}(\boldsymbol{q},\omega) \equiv \hat{\chi}(\boldsymbol{q},\omega) &= \frac{\boldsymbol{q}^2}{\omega^2}\left(\hat{\chi}_\parallel(\boldsymbol{q},\omega) - \frac{n}{m^*}\right)\\ \hat{\chi}_{0j}(\boldsymbol{q},\omega) = \hat{\chi}_{j0}(\boldsymbol{q},\omega) &= \frac{q_j}{\omega}\,\hat{\chi}_\parallel(\boldsymbol{q},\omega) \quad .\end{aligned} \tag{3.460}$$

Note that at $T = 0$ we have

$$\hat{\chi}(\boldsymbol{q},\omega) = \frac{1}{\hbar A}\sum_j\left[\frac{\langle\,0\,|\,\rho_{-\boldsymbol{q}}\,|\,j\,\rangle\,\langle\,j\,|\,\rho_{\boldsymbol{q}}\,|\,0\,\rangle}{\omega+\omega_j+i\epsilon} - \frac{\langle\,0\,|\,\rho_{\boldsymbol{q}}\,|\,j\,\rangle\,\langle\,j\,|\,\rho_{-\boldsymbol{q}}\,|\,0\,\rangle}{\omega-\omega_j+i\epsilon}\right] \tag{3.461}$$

where $\rho_{\boldsymbol{q}} = \sum_{i=1}^{N} e^{-i\boldsymbol{q}\cdot\boldsymbol{x}_i}$ is the Fourier component of the density. If the (normalized) SMA state,

$$|\,\boldsymbol{q}\,\rangle \equiv \big[Ns(\boldsymbol{q})\big]^{-1/2}\rho_{\boldsymbol{q}}\,|\,0\,\rangle \tag{3.462}$$

is an eigenstate, this means that $|\boldsymbol{q}\rangle$ saturates all the oscillator strength at this wavevector, in which case $\hat\chi(\boldsymbol{q},\omega) = \hat\chi_{\text{SMA}}(\boldsymbol{q},\omega)$, with

$$\hat\chi_{\text{SMA}}(\boldsymbol{q},\omega) = \frac{n}{\hbar}\,s(\boldsymbol{q})\left\{\frac{1}{\omega+\omega(\boldsymbol{q})+i\epsilon} - \frac{1}{\omega-\omega(\boldsymbol{q})+i\epsilon}\right\} = \frac{n\boldsymbol{q}^2/m}{\omega^2(\boldsymbol{q})-(\omega+i\epsilon)^2} \quad , \tag{3.463}$$

where $\omega(\boldsymbol{q}) = \langle\boldsymbol{q}|H|\boldsymbol{q}\rangle - E_0$ is the SMA energy.

Relation between charged and neutral system response

If we endow each of our neutral particles with a charge $(-e)$, then

$$\hat K_{\mu\nu}(\boldsymbol{q},\omega) = \frac{4\pi e^2}{c^2}\left(\frac{n}{m^*}\,\delta_{\mu\nu}(1-\delta_{\mu0}) - \hat\chi_{\mu\nu}(\boldsymbol{q},\omega)\right) \quad . \tag{3.464}$$

Thus,

$$\begin{aligned}
\hat K_{00}(\boldsymbol{q},\omega) &= \frac{4\pi e^2}{c^2}\,\frac{\boldsymbol{q}^2}{\omega^2}\left(\frac{n}{m^*} - \hat\chi_\parallel(\boldsymbol{q},\omega)\right) = -\frac{4\pi e^2}{c^2}\,\hat\chi(\boldsymbol{q},\omega) \\
\hat K_{0i}(\boldsymbol{q},\omega) = \hat K_{i0}(\boldsymbol{q},\omega) &= \frac{4\pi e^2}{c^2}\,\frac{q_i}{\omega}\left(\frac{n}{m^*} - \hat\chi_\parallel(\boldsymbol{q},\omega)\right) \\
\hat K_\parallel(\boldsymbol{q},\omega) &= \frac{4\pi e^2}{c^2}\left(\frac{n}{m^*} - \hat\chi_\parallel(\boldsymbol{q},\omega)\right) = \frac{\omega^2}{\boldsymbol{q}^2}\,\hat K_{00}(\boldsymbol{q},\omega) \\
\hat K_\perp(\boldsymbol{q},\omega) &= \frac{4\pi e^2}{c^2}\left(\frac{n}{m^*} - \hat\chi_\perp(\boldsymbol{q},\omega)\right) \quad .
\end{aligned} \tag{3.465}$$

3.8.4 Meissner effect and superfluid density

Suppose the electric field is $\boldsymbol{\mathcal{E}}$ and the magnetic field is $\boldsymbol{\mathcal{B}} = \boldsymbol{\nabla}\times\boldsymbol{\mathcal{A}}$ in three space dimensions. In $d = 2$ space dimensions, the magnetic field is a scalar $\mathcal{B} = \hat{\boldsymbol{z}}\cdot\boldsymbol{\nabla}\times\boldsymbol{\mathcal{A}}$, but we can also for convenience define $\boldsymbol{\mathcal{B}} = \mathcal{B}\hat{\boldsymbol{z}}$. We may work in the $\mathcal{A}^0 = 0$ gauge, in which case $\boldsymbol{\mathcal{E}} = -c^{-1}\partial_t\boldsymbol{\mathcal{A}}$, which entails $\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{A}} = 0$ at any finite frequency. We now have

$$\begin{aligned}
\boldsymbol{\nabla}\times\boldsymbol{\mathcal{A}} &= \frac{4\pi}{c}\,\boldsymbol{J} + \frac{1}{c}\,\frac{\partial\boldsymbol{\mathcal{E}}}{\partial t} = \boldsymbol{\nabla}(\boldsymbol{\nabla}\cdot\boldsymbol{\mathcal{A}}) - \nabla^2\boldsymbol{\mathcal{A}} \\
&= -\hat K_\perp(-i\boldsymbol{\nabla}, i\partial_t)\,\boldsymbol{\mathcal{A}} - \frac{1}{c^2}\,\frac{\partial^2\boldsymbol{\mathcal{A}}}{\partial t^2} \quad ,
\end{aligned} \tag{3.466}$$

which is valid for $d = 3$, but for $d = 2$ we must replace $\hat{K}_\perp \to \hat{K}_\perp / d_z$ where d_z is the thickness of the sample in the $\hat{\boldsymbol{z}}$ direction. One has that the units of $\hat{K}_{ij}(\boldsymbol{q}, \omega)$ are L^{1-d}. Thus

$$\nabla^2 \boldsymbol{\mathcal{A}} - \frac{1}{c^2} \frac{\partial^2 \boldsymbol{\mathcal{A}}}{\partial t^2} = \hat{K}_\perp(-i\boldsymbol{\nabla}, i\partial_t)\, \boldsymbol{\mathcal{A}} \quad , \tag{3.467}$$

again with $\hat{K}_\perp$ replaced by $\hat{K}_\perp / d_z$ in two space dimensions. Setting $\omega = 0$ and taking the limit $\boldsymbol{q} \to 0$, we obtain the formula for the London penetration depth,

$$\lambda_{\mathsf{L}}^{-2} = \left\{ \begin{matrix} d_z^{-1} \\ 1 \end{matrix} \right\} \times \lim_{\boldsymbol{q} \to 0} \hat{K}_\perp(\boldsymbol{q}, 0) \quad . \tag{3.468}$$

For a purely two-dimensional system, the distance between consecutive layers is $d_z = \infty$, and we have $\lambda_{\mathsf{L}} = \infty$, which says that a purely two-dimensional system cannot screen a three-dimensional electromagnetic field. In three-dimensions, the *superfluid density* is defined by

$$n_{\mathsf{s}} \equiv \frac{m^* c^2}{4\pi e^2 \lambda_{\mathsf{L}}^2} = n - m^* \lim_{\boldsymbol{q} \to 0} \hat{\chi}_\perp(\boldsymbol{q}, 0) \quad . \tag{3.469}$$

In the three-dimensional ideal Bose gas, for example, one finds

$$\hat{\chi}_{ij}(\boldsymbol{q} \to 0, 0) = \frac{n_0}{m^*}\, \hat{q}_i\, \hat{q}_j + \frac{n - n_0}{m^*}\, \delta_{ij} \quad , \tag{3.470}$$

where $n_0(T)$ is the number density of condensed bosons and $n' \equiv n - n_0$ is that of uncondensed bosons. Thus $\hat{\chi}_\parallel(\boldsymbol{q} \to 0, 0) = n/m^*$ and $\hat{\chi}_\perp(\boldsymbol{q} \to 0, 0) = n'/m^*$. The superfluid number density is $n_{\mathsf{s}}(T) = n_0(T)$. In fact, as Landau first showed, an ideal Bose gas is in fact *not* a superfluid because its excitation spectrum, which follows the ballistic dispersion $\omega(\boldsymbol{q}) = \hbar \boldsymbol{q}^2 / 2m$ is too 'soft', and any nonzero superflow is unstable to decay into single particle excitations.

Chapter 4

Beyond Laughlin

4.1 Landau level mixing

4.2 Chiral Luttinger Liquid Theory of FQH Edge States

4.3 Skyrmions at $\nu = 1$

4.4 Mesoscale Structure in Higher Landau Levels

4.5 Multilayers and QH Ferromagnetism

4.6 Coupled Wire Constructions

4.7 The Half-Filled Landau Level

4.8 The Moore-Read State and Nonabelions

4.9 Read-Rezayi States and Parafermions