Solid State Physics

Chetan Nayak

Contents

Chapter 1

What is Condensed Matter Physics?

1.1 Length, time, energy scales

We will be concerned with:

- $\omega, T \ll 1eV$

- $|x_i - x_j|, \frac{1}{q} \gg 1\mathring{A}$

as compared to energies in the MeV for nuclear matter, and GeV or even TeV, in particle physics.

The properties of matter at these scales is determined by the behavior of collections of many ($\sim 10^{23}$) atoms.

In general, we will be concerned with scales much smaller than those at which gravity becomes very important, which is the domain of astrophysics and cosmology.

1.2   Microscopic Equations vs. States of Matter, Phase Transitions, Critical Points

Systems containing many particles exhibit properties which are special to such systems. Many of these properties are fairly insensitive to the details at length scales shorter than $1\mathring{A}$ and energy scales higher than $1eV$ – which are quite adequately described by the equations of non-relativistic quantum mechanics. Such properties are *emergent*. For example, precisely the same microscopic equations of motion – Newton's equations – can describe two different systems of 10^{23} H_2O molecules.

$$m\frac{d^2\vec{x}_i}{dt^2} = -\sum_{j \neq i}\vec{\nabla}_i V\left(\vec{x}_i - \vec{x}_j\right) \tag{1.1}$$

Or, perhaps, the Schrödinger equation:

$$\left(-\frac{\hbar^2}{2m}\sum_i\nabla_i^2 + \sum_{i,j}V\left(\vec{x}_i - \vec{x}_j\right)\right)\psi\left(\vec{x}_1,\ldots,\vec{x}_N\right) = E\,\psi\left(\vec{x}_1,\ldots,\vec{x}_N\right) \tag{1.2}$$

However, one of these systems might be water and the other ice, in which case the properties of the two systems are completely different, and the similarity between their microscopic descriptions is of no practical consequence. As this example shows, many-particle systems exhibit various *phases* – such as ice and water – which are not, for the most part, usefully described by the microscopic equations. Instead, new low-energy, long-wavelength physics *emerges* as a result of the interactions among large numbers of particles. Different phases are separated by *phase transitions*, at which the low-energy, long-wavelength description becomes *non-analytic* and exhibits singularities. In the above example, this occurs at the freezing point of water, where its *entropy* jumps discontinuously.

1.3 Broken Symmetries

As we will see, different phases of matter are distinguished on the basis of symmetry. The microscopic equations are often highly symmetrical – for instance, Newton's laws are translationally and rotationally invariant – but a given phase may exhibit much less symmetry. Water exhibits the full translational and rotational symmetry of of Newton's laws; ice, however, is only invariant under the discrete translational and rotational group of its crystalline lattice. We say that the translational and rotational symmetries of the microscopic equations have been *spontaneously* broken.

1.4 Experimental probes: X-ray scattering, neutron scattering, NMR, thermodynamic, transport

There are various experimental probes which can allow an experimentalist to determine in what phase a system is and to determine its quantitative properties:

- Scattering: send neutrons or X-rays into the system with prescribed energy, momentum and measure the energy, momentum of the outgoing neutrons or X-rays.

- NMR: apply a static magnetic field, B, and measure the absorption and emission by the system of magnetic radiation at frequencies of the order of $\omega_c = geB/m$. Essentially the scattering of magnetic radiation at low frequency by a system in a uniform B field.

- Thermodynamics: measure the response of macroscopic variables such as the energy and volume to variations of the temperature, pressure, etc.

- Transport: set up a potential or thermal gradient, $\nabla\varphi$, ∇T and measure the electrical or heat current $\vec{j}$, $\vec{j}_Q$. The gradients $\nabla\varphi$, ∇T can be held constant or made to oscillate at finite frequency.

1.5 The Solid State: metals, insulators, magnets, superconductors

In the solid state, translational and rotational symmetries are broken by the arrangement of the positive ions. It is precisely as a result of these broken symmetries that solids are *solid*, i.e. that they are rigid. It is energetically favorable to break the symmetry in the same way in different parts of the system. Hence, the system resists attempts to create regions where the residual translational and rotational symmetry groups are different from those in the bulk of the system. The broken symmetry can be detected using X-ray or neutron scattering: the X-rays or neutrons are scattered by the ions; if the ions form a lattice, the X-rays or neutrons are scattered coherently, forming a diffraction pattern with peaks. In a crystalline solid, discrete subgroups of the translational and rotational groups are preserved. For instance, in a cubic lattice, rotations by $\pi/2$ about any of the crystal axes are symmetries of the lattice (as well as all rotations generated by products of these). Translations by one lattice spacing along a crystal axis generate the discrete group of translations.

In this course, we will be focussing on crystalline solids. Some examples of non-crystalline solids, such as plastics and glasses will be discussed below. Crystalline solids fall into three general categories: *metals*, *insulators*, and *superconductors*. In addition, all three of these phases can be further subdivided into various magnetic phases. Metals are characterized by a non-zero conductivity at $T = 0$. Insulators have vanishing conductivity at $T = 0$. Superconductors have infinite conductivity for

$T < T_c$ and, furthermore, exhibit the Meissner effect: they expel magnetic fields.

In a magnetic material, the electron spins can order, thereby breaking the spin-rotational invariance. In a ferromagnet, all of the spins line up in the same direction, thereby breaking the spin-rotational invariance to the subgroup of rotations about this direction while preserving the discrete translational symmetry of the lattice. (This can occur in a metal, an insulator, or a superconductor.) In an antiferromagnet, neighboring spins are oppositely directed, thereby breaking spin-rotational invariance to the subgroup of rotations about the preferred direction *and* breaking the lattice translational symmetry to the subgroup of translations by an even number of lattice sites.

Recently, new states of matter – the fractional quantum Hall states – have been discovered in effectively two-dimensional systems in a strong magnetic field at very low T. Tomorrow's experiments will undoubtedly uncover new phases of matter.

1.6 Other phases: liquid crystals, quasicrystals, polymers, glasses

The liquid – with full translational and rotational symmetry – and the solid – which only preserves a discrete subgroup – are but two examples of possible realizations of translational symmetry. In a *liquid crystalline phase*, translational and rotational symmetry is broken to a combination of discrete and continuous subgroups. For instance, a *nematic* liquid crystal is made up of elementary units which are line segments. In the nematic phase, these line segments point, on average, in the same direction, but their positional distribution is as in a liquid. Hence, a nematic phase breaks rotational invariance to the subgroup of rotations about the preferred direction and preserves the full translational invariance. Nematics are used in LCD displays.

In a *smectic* phase, on the other hand, the line segments arrange themselves into layers, thereby partially breaking the translational symmetry so that discrete translations perpendicular to the layers and continuous translations along the layers remain unbroken. In the *smectic-A* phase, the preferred orientational direction is the same as the direction perpendicular to the layers; in the *smectic-C phase*, these directions are different. In a hexatic phase, a two-dimensional system has broken orientational order, but unbroken translational order; locally, it looks like a triangular lattice. A *quasicrystal* has rotational symmetry which is broken to a 5-fold discrete subgroup. Translational order is completely broken (locally, it has discrete translational order). *Polymers* are extremely long molecules. They can exist in solution or a chemical reaction can take place which cross-links them, thereby forming a gel. A *glass* is a rigid, 'random' arrangement of atoms. Glasses are somewhat like 'snapshots' of liquids, and are probably non-equilibrium phases, in a sense.

Chapter 2

Review of Quantum Mechanics

2.1 States and Operators

A quantum mechanical system is defined by a Hilbert space, $\mathcal{H}$, whose vectors, $\left|\psi\right\rangle$ are associated with the states of the system. A state of the system is represented by the set of vectors $e^{i\alpha}\left|\psi\right\rangle$. There are linear operators, $\mathcal{O}_i$ which act on this Hilbert space. These operators correspond to physical observables. Finally, there is an inner product, which assigns a complex number, $\left\langle\chi|\psi\right\rangle$, to any pair of states, $\left|\psi\right\rangle$, $\left|\chi\right\rangle$. A state vector, $\left|\psi\right\rangle$ gives a complete description of a system through the expectation values, $\left\langle\psi|\mathcal{O}_i|\psi\right\rangle$ (assuming that $\left|\psi\right\rangle$ is normalized so that $\left\langle\psi|\psi\right\rangle = 1$), which would be the average values of the corresponding physical observables if we could measure them on an infinite collection of identical systems each in the state $\left|\psi\right\rangle$.

The adjoint, $\mathcal{O}^\dagger$, of an operator is defined according to

$$\left\langle\chi\right|\left(\mathcal{O}\left|\psi\right\rangle\right) = \left(\left\langle\chi\right|\mathcal{O}^\dagger\right)\left|\psi\right\rangle \tag{2.1}$$

In other words, the inner product between $\left|\chi\right\rangle$ and $\mathcal{O}\left|\psi\right\rangle$ is the same as that between $\mathcal{O}^\dagger\left|\chi\right\rangle$ and $\left|\psi\right\rangle$. An Hermitian operator satisfies

$$\mathcal{O} = \mathcal{O}^\dagger \tag{2.2}$$

while a unitary operator satisfies

$$\mathcal{O}\mathcal{O}^\dagger = \mathcal{O}^\dagger\mathcal{O} = 1 \tag{2.3}$$

If $\mathcal{O}$ is Hermitian, then

$$e^{i\mathcal{O}} \tag{2.4}$$

is unitary. Given an Hermitian operator, $\mathcal{O}$, its eigenstates are orthogonal,

$$\left\langle \lambda' \middle| \mathcal{O} \middle| \lambda \right\rangle = \lambda \left\langle \lambda' \middle| \lambda \right\rangle = \lambda' \left\langle \lambda' \middle| \lambda \right\rangle \tag{2.5}$$

For $\lambda \neq \lambda'$,

$$\left\langle \lambda' \middle| \lambda \right\rangle = 0 \tag{2.6}$$

If there are n states with the same eigenvalue, then, within the subspace spanned by these states, we can pick a set of n mutually orthogonal states. Hence, we can use the eigenstates $\left| \lambda \right\rangle$ as a basis for Hilbert space. Any state $\left| \psi \right\rangle$ can be expanded in the basis given by the eigenstates of $\mathcal{O}$:

$$\left| \psi \right\rangle = \sum_\lambda c_\lambda \left| \lambda \right\rangle \tag{2.7}$$

with

$$c_\lambda = \left\langle \lambda \middle| \psi \right\rangle \tag{2.8}$$

A particularly important operator is the Hamiltonian, or the total energy, which we will denote by H. Schrödinger's equation tells us that H determines how a state of the system will evolve in time.

$$i\hbar \frac{\partial}{\partial t} \left| \psi \right\rangle = H \left| \psi \right\rangle \tag{2.9}$$

If the Hamiltonian is independent of time, then we can define energy eigenstates,

$$H \left| E \right\rangle = E \left| E \right\rangle \tag{2.10}$$

which evolve in time according to:

$$\left|E(t)\right\rangle = e^{-i\frac{Et}{\hbar}}\left|E(0)\right\rangle \tag{2.11}$$

An arbitrary state can be expanded in the basis of energy eigenstates:

$$\left|\psi\right\rangle = \sum_i c_i \left|E_i\right\rangle \tag{2.12}$$

It will evolve according to:

$$\left|\psi(t)\right\rangle = \sum_j c_j e^{-i\frac{E_j t}{\hbar}}\left|E_j\right\rangle \tag{2.13}$$

For example, consider a particle in $1D$. The Hilbert space consists of all continuous complex-valued functions, $\psi(x)$. The position operator, $\hat{x}$, and momentum operator, $\hat{p}$ are defined by:

$$\begin{aligned}
\hat{x} \cdot \psi(x) &\equiv x\,\psi(x) \\
\hat{p} \cdot \psi(x) &\equiv -i\hbar \frac{\partial}{\partial x}\,\psi(x)
\end{aligned} \tag{2.14}$$

The position eigenfunctions,

$$x\,\delta(x-a) = a\,\delta(x-a) \tag{2.15}$$

are Dirac delta functions, which are not continuous functions, but can be defined as the limit of continuous functions:

$$\delta(x) = \lim_{a\to 0}\frac{1}{a\sqrt{\pi}}\,e^{-\frac{x^2}{a^2}} \tag{2.16}$$

The momentum eigenfunctions are plane waves:

$$-i\hbar\frac{\partial}{\partial x}\,e^{ikx} = \hbar k\,e^{ikx} \tag{2.17}$$

Expanding a state in the basis of momentum eigenstates is the same as taking its Fourier transform:

$$\psi(x) = \int_{-\infty}^{\infty} dk\,\tilde{\psi}(k)\,\frac{1}{\sqrt{2\pi}}e^{ikx} \tag{2.18}$$

where the Fourier coefficients are given by:

$$\tilde{\psi}(k) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dx \, \psi(x) \, e^{-ikx} \tag{2.19}$$

If the particle is free,

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \tag{2.20}$$

then momentum eigenstates are also energy eigenstates:

$$\hat{H} e^{ikx} = \frac{\hbar^2 k^2}{2m} e^{ikx} \tag{2.21}$$

If a particle is in a Gaussian wavepacket at the origin at time $t = 0$,

$$\psi(x, 0) = \frac{1}{a\sqrt{\pi}} e^{-\frac{x^2}{a^2}} \tag{2.22}$$

Then, at time t, it will be in the state:

$$\psi(x, t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} dk \, \frac{a}{\sqrt{\pi}} e^{-i\frac{\hbar k^2 t}{2m}} e^{-\frac{1}{2}k^2 a^2} e^{ikx} \tag{2.23}$$

2.2 Density and Current

Multiplying the free-particle Schrödinger equation by ψ^*,

$$\psi^* i\hbar \frac{\partial}{\partial t} \psi = -\frac{\hbar^2}{2m} \psi^* \frac{\partial^2}{\nabla^2} \psi \tag{2.24}$$

and subtracting the complex conjugate of this equation, we find

$$\frac{\partial}{\partial t} \left(\psi^* \psi \right) = \frac{i\hbar}{2m} \vec{\nabla} \cdot \left(\psi^* \vec{\nabla} \psi - \left(\vec{\nabla} \psi^* \right) \psi \right) \tag{2.25}$$

This is in the form of a continuity equation,

$$\frac{\partial \rho}{\partial t} = \vec{\nabla} \cdot \vec{j} \tag{2.26}$$

The density and current are given by:

$$\rho = \psi^* \psi$$

$$\vec{j} = \frac{i\hbar}{2m} \left(\psi^* \vec{\nabla} \psi - \left(\vec{\nabla} \psi^* \right) \psi \right) \tag{2.27}$$

The current carried by a plane-wave state is:

$$\vec{j} = \frac{\hbar}{2m} \vec{k} \frac{1}{(2\pi)^3} \tag{2.28}$$

2.3 δ-function scatterer

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + V \delta(x) \tag{2.29}$$

$$\psi(x) = \begin{cases} e^{ikx} + R e^{-ikx} & \text{if } x < 0 \\ T e^{ikx} & \text{if } x > 0 \end{cases} \tag{2.30}$$

$$\begin{aligned} T &= \frac{1}{1 - \frac{mV}{\hbar^2 k} i} \\ R &= \frac{\frac{mV}{\hbar^2 k} i}{1 - \frac{mV}{\hbar^2 k} i} \end{aligned} \tag{2.31}$$

There is a *bound state* at:

$$ik = \frac{mV}{\hbar^2} \tag{2.32}$$

2.4 Particle in a Box

Particle in a $1D$ region of length L:

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} \tag{2.33}$$

$$\psi(x) = A e^{ikx} + B e^{-ikx} \tag{2.34}$$

has energy $E = \hbar^2 k^2 / 2m$. $\psi(0) = \psi(L) = 0$. Therefore,

$$\psi(x) = A \sin \left(\frac{n\pi}{L} x \right) \tag{2.35}$$

for integer n. Allowed energies

$$E_n = \frac{\hbar^2 \pi^2 n^2}{2mL^2} \tag{2.36}$$

In a $3D$ box of side L, the energy eigenfunctions are:

$$\psi(x) = A \sin\left(\frac{n_x \pi}{L} x\right) \sin\left(\frac{n_y \pi}{L} y\right) \sin\left(\frac{n_z \pi}{L} z\right) \tag{2.37}$$

and the allowed energies are:

$$E_n = \frac{\hbar^2 \pi^2}{2mL^2} \left(n_x^2 + n_y^2 + n_z^2\right) \tag{2.38}$$

2.5 Harmonic Oscillator

$$H = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} k x^2 \tag{2.39}$$

Writing $\omega = \sqrt{k/m}$, $\tilde{p} = p/(km)^{1/4}$, $\tilde{x} = x(km)^{1/4}$,

$$H = \frac{1}{2} \omega \left(\tilde{p}^2 + \tilde{x}^2\right) \tag{2.40}$$

$$[\tilde{p}, \tilde{x}] = -i\hbar \tag{2.41}$$

Raising and lowering operators:

$$a = (\tilde{x} + i\tilde{p})/\sqrt{2\hbar}$$
$$a^\dagger = (\tilde{x} - i\tilde{p})/\sqrt{2\hbar}$$

$$\tag{2.42}$$

Hamiltonian and commutation relations:

$$\begin{aligned} H &= \hbar\omega \left(a^\dagger a + \frac{1}{2}\right) \\ [a, a^\dagger] &= 1 \end{aligned} \tag{2.43}$$

The commutation relations,

$$[H, a^\dagger] = \hbar\omega a^\dagger$$

$$[H, a] = -\hbar\omega a \tag{2.44}$$

imply that there is a ladder of states,

$$Ha^\dagger|E\rangle = (E + \hbar\omega) a^\dagger|E\rangle$$
$$Ha|E\rangle = (E - \hbar\omega) a|E\rangle \tag{2.45}$$

This ladder will continue down to negative energies (which it can't since the Hamiltonian is manifestly positive definite) unless there is an $E_0 \geq 0$ such that

$$a|E_0\rangle = 0 \tag{2.46}$$

Such a state has $E_0 = \hbar\omega/2$.

We label the states by their $a^\dagger a$ eigenvalues. We have a complete set of H eigenstates, $|n\rangle$, such that

$$H|n\rangle = \hbar\omega \left(n + \frac{1}{2} \right) |n\rangle \tag{2.47}$$

and $(a^\dagger)^n|0\rangle \propto |n\rangle$. To get the normalization, we write $a^\dagger|n\rangle = c_n|n+1\rangle$. Then,

$$\begin{aligned} |c_n|^2 &= \langle n|aa^\dagger|n\rangle \\ &= n + 1 \end{aligned} \tag{2.48}$$

Hence,

$$\begin{aligned} a^\dagger|n\rangle &= \sqrt{n+1}|n+1\rangle \\ a|n\rangle &= \sqrt{n}|n-1\rangle \end{aligned} \tag{2.49}$$

2.6 Double Well

$$H = -\frac{\hbar^2}{2m}\frac{\partial^2}{\partial x^2} + V(x) \tag{2.50}$$

where

$$V(x) = \begin{cases} \infty & \text{if } |x| > 2a + 2b \\ 0 & \text{if } b < |x| < a + b \\ V_0 & \text{if } |x| < b \end{cases}$$

Symmetrical solutions:

$$\psi(x) = \begin{cases} A \cos k'x & \text{if } |x| < b \\ \cos(k|x| - \phi) & \text{if } b < |x| < a + b \end{cases} \tag{2.51}$$

with

$$k' = \sqrt{k^2 - \frac{2mV_0}{\hbar^2}} \tag{2.52}$$

The allowed k's are determined by the condition that $\psi(a + b) = 0$:

$$\phi = \left(n + \frac{1}{2}\right)\pi - k(a + b) \tag{2.53}$$

the continuity of $\psi(x)$ at $|x| = b$:

$$A = \frac{\cos(kb - \phi)}{\cos k'b} \tag{2.54}$$

and the continuity of $\psi'(x)$ at $|x| = b$:

$$k \tan\left(\left(n + \frac{1}{2}\right)\pi - ka\right) = k' \tan k'b \tag{2.55}$$

If k' is imaginary, $\cos \to \cosh$ and $\tan \to i \tanh$ in the above equations.

Antisymmetrical solutions:

$$\psi(x) = \begin{cases} A \sin k'x & \text{if } |x| < b \\ \text{sgn}(x) \cos(k|x| - \phi) & \text{if } b < |x| < a + b \end{cases} \tag{2.56}$$

The allowed k's are now determined by

$$\phi = \left(n + \frac{1}{2}\right)\pi - k(a + b) \tag{2.57}$$

$$A = \frac{\cos(kb - \phi)}{\sin k'b} \tag{2.58}$$

$$k \tan\left(\left(n + \frac{1}{2}\right)\pi - ka\right) = -k' \cot k'b \qquad (2.59)$$

Suppose we have n wells? Sequences of eigenstates, classified according to their eigenvalues under translations between the wells.

2.7   Spin

The electron carries spin-1/2. The spin is described by a state in the Hilbert space:

$$\alpha|+\rangle \;\; + \;\; \beta|-\rangle \qquad (2.60)$$

spanned by the basis vectors $|\pm\rangle$. Spin operators:

$$
\begin{aligned}
s_x &= \frac{1}{2}\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\
s_y &= \frac{1}{2}\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\
s_z &= \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}
\end{aligned}
\qquad (2.61)
$$

Coupling to an external magnetic field:

$$H_{\text{int}} = -g\mu_B \vec{s} \cdot \vec{B} \qquad (2.62)$$

States of a spin in a magnetic field in the $\hat{\mathbf{z}}$ direction:

$$
\begin{aligned}
H|+\rangle &= -\frac{g}{2}\mu_B\,|+\rangle \\
H|-\rangle &= \frac{g}{2}\mu_B\,|-\rangle
\end{aligned}
\qquad (2.63)
$$

2.8   Many-Particle Hilbert Spaces: Bosons, Fermions

When we have a system with many particles, we must now specify the states of all of the particles. If we have two distinguishable particles whose Hilbert spaces are

spanned by the bases

$$\left| i, 1 \right\rangle \tag{2.64}$$

and

$$\left| \alpha, 2 \right\rangle \tag{2.65}$$

Then the two-particle Hilbert space is spanned by the set:

$$\left| i, 1; \alpha, 2 \right\rangle \equiv \left| i, 1 \right\rangle \otimes \left| \alpha, 2 \right\rangle \tag{2.66}$$

Suppose that the two single-particle Hilbert spaces are identical, e.g. the two particles are in the same box. Then the two-particle Hilbert space is:

$$\left| i, j \right\rangle \equiv \left| i, 1 \right\rangle \otimes \left| j, 2 \right\rangle \tag{2.67}$$

If the particles are identical, however, we must be more careful. $\left| i, j \right\rangle$ and $\left| j, i \right\rangle$ must be physically the same state, i.e.

$$\left| i, j \right\rangle = e^{i\alpha} \left| j, i \right\rangle \tag{2.68}$$

Applying this relation twice implies that

$$\left| i, j \right\rangle = e^{2i\alpha} \left| i, j \right\rangle \tag{2.69}$$

so $e^{i\alpha} = \pm 1$. The former corresponds to bosons, while the latter corresponds to fermions. The two-particle Hilbert spaces of bosons and fermions are respectively spanned by:

$$\left| i, j \right\rangle + \left| j, i \right\rangle \tag{2.70}$$

and

$$\left| i, j \right\rangle - \left| j, i \right\rangle \tag{2.71}$$

The n-particle Hilbert spaces of bosons and fermions are respectively spanned by:

$$\sum_{\pi} \left| i_{\pi(1)}, \ldots, i_{\pi(n)} \right\rangle \tag{2.72}$$

and

$$\sum_{\pi}(-1)^{\pi}\left|i_{\pi(1)},\ldots,i_{\pi(n)}\right\rangle \tag{2.73}$$

In position space, this means that a bosonic wavefunction must be completely symmetric:

$$\psi(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = \psi(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n) \tag{2.74}$$

while a fermionic wavefunction must be completely antisymmetric:

$$\psi(x_1,\ldots,x_i,\ldots,x_j,\ldots,x_n) = -\psi(x_1,\ldots,x_j,\ldots,x_i,\ldots,x_n) \tag{2.75}$$

Chapter 3

Review of Statistical Mechanics

3.1 Microcanonical, Canonical, Grand Canonical Ensembles

In statistical mechanics, we deal with a situation in which even the quantum state of the system is unknown. The expectation value of an observable must be averaged over:

$$\langle \mathcal{O} \rangle = \sum_i w_i \langle i \left| \mathcal{O} \right| i \rangle \tag{3.1}$$

where the states $|i\rangle$ form an orthonormal basis of $\mathcal{H}$ and w_i is the probability of being in state $|i\rangle$. The w_i's must satisfy $\sum w_i = 1$. The expectation value can be written in a basis-independent form:

$$\langle \mathcal{O} \rangle = Tr \left\{ \rho \mathcal{O} \right\} \tag{3.2}$$

where ρ is the density matrix. In the above example, $\rho = \sum_i w_i |i\rangle\langle i|$. The condition, $\sum w_i = 1$, i.e. that the probabilities add to 1, is:

$$Tr \left\{ \rho \right\} = 1 \tag{3.3}$$

We usually deal with one of three ensembles: the microcanonical emsemble, the canonical ensemble, or the grand canonical ensemble. In the microcanonical ensemble,

18

we assume that our system is isolated, so the energy is fixed to be E, but all states with energy E are taken with equal probability:

$$\rho = C\,\delta(H - E) \tag{3.4}$$

C is a normalization constant which is determined by (3.3). The entropy is given by,

$$S = -\ln C \tag{3.5}$$

In other words,

$$S(E) = \ln\Big(\#\text{ of states with energy } E\Big) \tag{3.6}$$

Inverse temperature, $\beta = 1/(k_B T)$:

$$\beta \equiv \left(\frac{\partial S}{\partial E}\right)_V \tag{3.7}$$

Pressure, P:

$$\frac{P}{k_B T} \equiv \left(\frac{\partial S}{\partial V}\right)_E \tag{3.8}$$

where V is the volume.

First law of thermodynamics:

$$dS = \frac{\partial S}{\partial E}\,dE + \frac{\partial S}{\partial V}\,dV \tag{3.9}$$

$$dE = k_B T\,dS - P\,dV \tag{3.10}$$

Free energy:

$$F = E - k_B T S \tag{3.11}$$

Differential relation:

$$dF = -k_B S\,dT - P\,dV \tag{3.12}$$

or,

$$S = -\frac{1}{k_B}\left(\frac{\partial F}{\partial T}\right)_V \tag{3.13}$$

$$P = -\left(\frac{\partial F}{\partial V}\right)_T \tag{3.14}$$

while

$$
\begin{aligned}
E &= F + k_B T S \\
&= F - T \left(\frac{\partial F}{\partial T}\right)_V \\
&= -T^2 \frac{\partial}{\partial T} \frac{F}{T}
\end{aligned}
\tag{3.15}
$$

In the canonical ensemble, we assume that our system is in contact with a heat reservoir so that the *temperature is constant*. Then,

$$\rho = C\,e^{-\beta H} \tag{3.16}$$

It is useful to drop the normalization constant, C, and work with an unnormalized density matrix so that we can define the partition function:

$$Z = Tr\{\rho\} \tag{3.17}$$

or,

$$Z = \sum_a e^{-\beta E_a} \tag{3.18}$$

The average energy is:

$$
\begin{aligned}
E &= \frac{1}{Z}\sum_a E_a\,e^{-\beta E_a} \\
&= -\frac{\partial}{\partial \beta}\ln Z \\
&= -k_B T^2 \frac{\partial}{\partial T}\ln Z
\end{aligned}
\tag{3.19}
$$

Hence,

$$F = -k_B T \ln Z \tag{3.20}$$

The chemical potential, μ, is defined by

$$\mu = \frac{\partial F}{\partial N} \tag{3.21}$$

where N is the particle number.

In the grand canonical ensemble, the system is in contact with a reservoir of heat *and* particles. Thus, the temperature and chemical potential are held fixed and

$$\rho = C \, e^{-\beta(H-\mu N)} \tag{3.22}$$

We can again work with an unnormalized density matrix and construct the grand canonical partition function:

$$\mathcal{Z} = \sum_{N,a} e^{-\beta(E_a-\mu N)} \tag{3.23}$$

The average number is:

$$N = -k_B T \, \frac{\partial}{\partial \mu} \, \ln \mathcal{Z} \tag{3.24}$$

while the average energy is:

$$E = -\frac{\partial}{\partial \beta} \, \ln \mathcal{Z} + \mu k_B T \, \frac{\partial}{\partial \mu} \, \ln \mathcal{Z} \tag{3.25}$$

3.2   Bose-Einstein and Planck Distributions

3.2.1   Bose-Einstein Statistics

For a system of free bosons, the partition function

$$Z = \sum_{E_a,N} e^{-\beta(E_a-\mu N)} \tag{3.26}$$

can be rewritten in terms of the *single-particle eigenstates* and the single-particle energies ϵ_i:

$$E_a = n_0 \, \epsilon_0 + n_1 \, \epsilon_1 + \dots \tag{3.27}$$

$$
\begin{aligned}
Z &= \sum_{\{n_i\}} e^{-\beta\left(\sum_i n_i\epsilon_i - \mu\sum_i n_i\right)} \\
&= \prod_i \left(\sum_{n_i} e^{-\beta(n_i\epsilon_i - \mu n_i)} \right)
\end{aligned}
$$

$$= \prod_i \frac{1}{1 - e^{-\beta(\epsilon_i - \mu)}} \tag{3.28}$$

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1} \tag{3.29}$$

The chemical potential is chosen so that

$$\begin{aligned} N &= \sum_i \langle n_i \rangle \\ &= \sum_i \frac{1}{e^{\beta(\epsilon_i - \mu)} - 1} \end{aligned} \tag{3.30}$$

The energy is given by

$$\begin{aligned} E &= \sum_i \langle n_i \rangle \, \epsilon_i \\ &= \sum_i \frac{\epsilon_i}{e^{\beta(\epsilon_i - \mu)} - 1} \end{aligned} \tag{3.31}$$

N is increased by increasing μ ($\mu \leq 0$ always). Bose-Einstein condensation occurs when

$$N > \sum_{i \neq 0} \langle n_i \rangle \tag{3.32}$$

In such a case, $\langle n_0 \rangle$ must become large. This occurs when $\mu = 0$.

3.2.2 The Planck Distribution

Suppose N is not fixed, but is arbitrary, e.g. the numbers of photons and neutrinos are not fixed. Then there is no Lagrange multiplier μ and

$$\langle n_i \rangle = \frac{1}{e^{\beta \epsilon_i} - 1} \tag{3.33}$$

Consider photons (two polarizations) in a cavity of side L with $\epsilon_k = \hbar \omega_k = \hbar c k$ and

$$k = \frac{2\pi}{L} \, (m_x, m_y, m_z) \tag{3.34}$$

$$E = 2 \sum_{m_x, m_y, m_z} \omega_{m_x, m_y, m_z} \left\langle n_{m_x, m_y, m_z} \right\rangle \tag{3.35}$$

We can take the thermodynamic limit, $L \to \infty$, and convert the sum into an integral. Since the allowed $\vec{k}$'s are $\frac{2\pi}{L}(m_x, m_y, m_z)$, the $\vec{k}$-space volume per allowed $\vec{k}$ is $(2\pi)^3/L^3$. Hence, we can take the infinite-volume limit by making the replacement:

$$
\begin{aligned}
\sum_k f(\vec{k}) &= \frac{1}{(\Delta \vec{k})^3} \sum_k f(\vec{k}) (\Delta \vec{k})^3 \\
&= \frac{L^3}{(2\pi)^3} \int d^3\vec{k}\, f(\vec{k})
\end{aligned}
\tag{3.36}
$$

Hence,

$$
\begin{aligned}
E &= 2V \int_0^{k_{\max}} \frac{d^3k}{(2\pi)^3} \frac{\hbar \omega_k}{e^{\beta \hbar \omega_k} - 1} \\
&= 2V \int_0^{k_{\max}} \frac{d^3k}{(2\pi)^3} \frac{\hbar c k}{e^{\beta \hbar c k} - 1} \\
&= \frac{V k_B^4}{\pi^2 (\hbar c)^3} T^4 \int_0^{\beta \hbar c k_{\max}} \frac{x^3\, dx}{e^x - 1}
\end{aligned}
\tag{3.37}
$$

For $\beta \hbar c k_{\max} \gg 1$,

$$
E = \frac{V k_B^4}{\pi^2 (\hbar c)^3} T^4 \int_0^\infty \frac{x^3\, dx}{e^x - 1}
\tag{3.38}
$$

and

$$
C_V = \frac{4V k_B^4}{\pi^2 (\hbar c)^3} T^3 \int_0^\infty \frac{x^3\, dx}{e^x - 1}
\tag{3.39}
$$

For $\beta \hbar c k_{\max} \ll 1$,

$$
E = \frac{V k_{\max}^3}{3\pi^2} k_B T
\tag{3.40}
$$

and

$$
C_V = \frac{V k_{\max}^3 k_B}{3\pi^2}
\tag{3.41}
$$

3.3 Fermi-Dirac Distribution

For a system of free fermions, the partition function

$$
Z = \sum_{E_a, N} e^{-\beta(E_a - \mu N)}
\tag{3.42}
$$

can again be rewritten in terms of the *single-particle eigenstates* and the single-particle energies ϵ_i:

$$E_a = n_0\,\epsilon_0 + n_1\,\epsilon_1 + \ldots \tag{3.43}$$

but now

$$n_i = 0, 1 \tag{3.44}$$

so that

$$
\begin{aligned}
Z &= \sum_{\{n_i\}} e^{-\beta\left(\sum_i n_i\epsilon_i - \mu\sum_i n_i\right)} \\
&= \prod_i \left(\sum_{n_i=0}^{1} e^{-\beta(n_i\epsilon_i - \mu n_i)}\right) \\
&= \prod_i \left(1 + e^{-\beta(\epsilon_i - \mu)}\right)
\end{aligned}
\tag{3.45}
$$

$$\langle n_i \rangle = \frac{1}{e^{\beta(\epsilon_i + \mu)} + 1} \tag{3.46}$$

The chemical potential is chosen so that

$$N = \sum_i \frac{1}{e^{\beta(\epsilon_i + \mu)} + 1} \tag{3.47}$$

The energy is given by

$$E = \sum_i \frac{\epsilon_i}{e^{\beta(\epsilon_i + \mu)} + 1} \tag{3.48}$$

3.4 Thermodynamics of the Free Fermion Gas

Free electron gas in a box of side L:

$$\epsilon_k = \frac{\hbar^2 k^2}{2m} \tag{3.49}$$

with

$$k = \frac{2\pi}{L}\,(m_x, m_y, m_z) \tag{3.50}$$

Then, taking into account the 2 spin states,

$$
\begin{aligned}
E &= 2 \sum_{m_x, m_y, m_z} \epsilon_{m_x, m_y, m_z} \left\langle n_{m_x, m_y, m_z} \right\rangle \\
&= 2V \int_0^{k_{\max}} \frac{d^3k}{(2\pi)^3} \frac{\frac{\hbar^2 k^2}{2m}}{e^{\beta\left(\frac{\hbar^2 k^2}{2m} - \mu\right)} + 1}
\end{aligned}
$$
(3.51)

$$
N = 2V \int_0^{k_{\max}} \frac{d^3k}{(2\pi)^3} \frac{1}{e^{\beta\left(\frac{\hbar^2 k^2}{2m} - \mu\right)} + 1}
$$
(3.52)

At $T = 0$,

$$
\frac{1}{e^{\beta\left(\frac{\hbar^2 k^2}{2m} - \mu\right)} + 1} = \theta\left(\mu - \frac{\hbar^2 k^2}{2m}\right)
$$
(3.53)

All states with energies less than μ are filled; all states with higher energies are empty.

We write

$$
k_F = \frac{\sqrt{2m\mu_{T=0}}}{\hbar} \ , \qquad \epsilon_F = \mu_{T=0}
$$
(3.54)

$$
\frac{N}{V} = 2 \int_0^{k_F} \frac{d^3k}{(2\pi)^3} = \frac{k_F^3}{3\pi^2}
$$
(3.55)

$$
\begin{aligned}
\frac{E}{V} &= 2 \int_0^{k_F} \frac{d^3k}{(2\pi)^3} \frac{\hbar^2 k^2}{2m} \\
&= \frac{1}{\pi^2} \frac{\hbar^2 k_F^5}{10m} \\
&= \frac{3}{5} \frac{N}{V} \epsilon_F
\end{aligned}
$$
(3.56)

$$
2 \int \frac{d^3k}{(2\pi)^3} = \frac{m^{\frac{3}{2}} 2^{\frac{1}{2}}}{\pi^2 \hbar^3} \int d\epsilon \, \epsilon^{\frac{1}{2}}
$$
(3.57)

For $k_B T \ll e_F$,

$$
\begin{aligned}
\frac{N}{V} &= \frac{m^{\frac{3}{2}} 2^{\frac{1}{2}}}{\pi^2 \hbar^3} \int_0^\infty d\epsilon \, \epsilon^{\frac{1}{2}} \frac{1}{e^{\beta(\epsilon - \mu)} + 1} \\
&= \frac{m^{\frac{3}{2}} 2^{\frac{1}{2}}}{\pi^2 \hbar^3} \int_0^\mu d\epsilon \, \epsilon^{\frac{1}{2}} + \frac{m^{\frac{3}{2}} 2^{\frac{1}{2}}}{\pi^2 \hbar^3} \int_0^\mu d\epsilon \, \epsilon^{\frac{1}{2}} \left(\frac{1}{e^{\beta(\epsilon - \mu)} + 1} - 1\right) + \frac{m^{\frac{3}{2}} 2^{\frac{1}{2}}}{\pi^2 \hbar^3} \int_\mu^\infty d\epsilon \, \epsilon^{\frac{1}{2}} \frac{1}{e^{\beta(\epsilon - \mu)} + 1}
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(2m)^{\frac{3}{2}}}{3\pi^2\hbar^3}\mu^{\frac{3}{2}} - \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\mu d\epsilon\, \epsilon^{\frac{1}{2}}\frac{1}{e^{-\beta(\epsilon-\mu)}+1} + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_\mu^\infty d\epsilon\, \epsilon^{\frac{1}{2}}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \frac{(2m)^{\frac{3}{2}}}{3\pi^2\hbar^3}\mu^{\frac{3}{2}} + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\infty \frac{k_B T\, dx}{e^x+1}\left((\mu+k_B Tx)^{\frac{1}{2}} - (\mu-k_B Tx)^{\frac{1}{2}}\right) + O\left(e^{-\beta\mu}\right) \\
&= \frac{(2m)^{\frac{3}{2}}}{3\pi^2\hbar^3}\mu^{\frac{3}{2}} + \frac{(2m)^{\frac{3}{2}}}{\pi^2\hbar^3}\sum_{n=1}^\infty (k_B T)^{2n}\left[\mu^{\frac{3}{2}-2n}\frac{1}{(2n-1)!}\frac{\Gamma\left(\frac{3}{2}\right)}{\Gamma\left(\frac{5}{2}-2n\right)}\int_0^\infty dx\,\frac{x^{2n-1}}{e^x+1}\right] \\
&= \frac{(2m)^{\frac{3}{2}}}{3\pi^2\hbar^3}\mu^{\frac{3}{2}}\left[1 + \frac{3}{2}\left(\frac{k_B T}{\mu}\right)^2 I_1 + O\left(T^4\right)\right]
\end{aligned}
\tag{3.58}
$$

with

$$
I_k = \int_0^\infty dx\,\frac{x^k}{e^x+1}
\tag{3.59}
$$

We will only need

$$
I_1 = \frac{\pi^2}{12}
\tag{3.60}
$$

Hence,

$$
(\epsilon_F)^{\frac{3}{2}} = \mu^{\frac{3}{2}}\left[1 + \frac{3}{2}\left(\frac{k_B T}{\mu}\right)^2 I_1 + O\left(T^4\right)\right]
\tag{3.61}
$$

To lowest order in T, this gives:

$$
\begin{aligned}
\mu &= \epsilon_F\left(1 - \left(\frac{k_B T}{\epsilon_F}\right)^2 I_1 + O(T^4)\right) \\
&= \epsilon_F\left(1 - \frac{\pi^2}{12}\left(\frac{k_B T}{\epsilon_F}\right)^2 + O(T^4)\right)
\end{aligned}
\tag{3.62}
$$

$$
\begin{aligned}
\frac{E}{V} &= \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\infty d\epsilon\, \epsilon^{\frac{3}{2}}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\mu d\epsilon\, \epsilon^{\frac{3}{2}} + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\mu d\epsilon\, \epsilon^{\frac{3}{2}}\left(\frac{1}{e^{\beta(\epsilon-\mu)}+1} - 1\right) + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_\mu^\infty d\epsilon\, \epsilon^{\frac{3}{2}}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \frac{(2m)^{\frac{3}{2}}}{5\pi^2\hbar^3}\mu^{\frac{5}{2}} - \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\mu d\epsilon\, \epsilon^{\frac{3}{2}}\frac{1}{e^{-\beta(\epsilon-\mu)}+1} + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_\mu^\infty d\epsilon\, \epsilon^{\frac{3}{2}}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \frac{(2m)^{\frac{3}{2}}}{5\pi^2\hbar^3}\mu^{\frac{5}{2}} + \frac{m^{\frac{3}{2}}2^{\frac{1}{2}}}{\pi^2\hbar^3}\int_0^\infty \frac{k_B T\, dx}{e^x+1}\left((\mu+k_B Tx)^{\frac{3}{2}} - (\mu-k_B Tx)^{\frac{3}{2}}\right) + O\left(e^{-\beta\mu}\right) \\
&= \frac{(2m)^{\frac{3}{2}}}{5\pi^2\hbar^3}\mu^{\frac{5}{2}} + \frac{(2m)^{\frac{3}{2}}}{\pi^2\hbar^3}\sum_{n=1}^\infty (k_B T)^{2n}\left[\mu^{\frac{5}{2}-2n}\frac{1}{(2n-1)!}\frac{\Gamma\left(\frac{5}{2}\right)}{\Gamma\left(\frac{7}{2}-2n\right)}\int_0^\infty dx\,\frac{x^{2n-1}}{e^x+1}\right]
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{(2m)^{\frac{3}{2}}}{5\pi^2\hbar^3}\mu^{\frac{5}{2}}\left[1 + \frac{15}{2}\left(\frac{k_BT}{\mu}\right)^2 I_1 + O(T^4)\right] \\
&= \frac{3}{5}\frac{N}{V}\epsilon_F\left(1 + \frac{5\pi^2}{12}\left(\frac{k_BT}{\epsilon_F}\right)^2 + O(T^4)\right)
\end{aligned}
\tag{3.63}
$$

Hence, the specific heat of a gas of free fermions is:

$$
C_V = \frac{\pi^2}{2}Nk_B\frac{k_BT}{\epsilon_F}
\tag{3.64}
$$

Note that this can be written in the more general form:

$$
C_V = (\text{const.}) \cdot k_B \cdot g\left(\epsilon_F\right) k_BT
\tag{3.65}
$$

The number of electrons which are thermally excited above the ground state is $\sim$ $g\left(\epsilon_F\right) k_BT$; each such electron contributes energy $\sim k_BT$ and, hence, gives a specific heat contribution of k_B. Electrons give such a contribution to the specific heat of a metal.

3.5 Ising Model, Mean Field Theory, Phases

Consider a model of spins on a lattice in a magnetic field:

$$
H = -g\mu_B B\sum_i S_i^z \equiv 2h\sum_i S_i^z
\tag{3.66}
$$

with $S_i^z = \pm 1/2$. The partition function for such a system is:

$$
Z = \left(2\cosh\frac{h}{k_BT}\right)^N
\tag{3.67}
$$

The average magnetization is:

$$
S_i^z = \frac{1}{2}\tanh\frac{h}{k_BT}
\tag{3.68}
$$

The susceptibility, χ, is defined by

$$
\chi = \left(\frac{\partial}{\partial h}\sum_i S_i^z\right)_{h=0}
\tag{3.69}
$$

For free spins on a lattice,

$$\chi = \frac{1}{2}\, N\, \frac{1}{k_B T} \tag{3.70}$$

A susceptibility which is inversely proportional to temperature is called a *Curie* susceptibility. In problem set 3, you will show that the susceptibility is much smaller for a system of electrons.

Now consider a model of spins on a lattice such that each spin interacts with its neighbors according to:

$$H = -\frac{1}{2}\sum_{\langle i,j \rangle} J S_i^z S_j^z \tag{3.71}$$

This Hamiltonian has a symmetry

$$S_i^z \to -S_i^z \tag{3.72}$$

For $k_B T \gg J$, the interaction between the spins will not be important and the susceptibility will be of the Curie form. For $k_B T < J$, however, the behavior will be much different. We can understand this qualitatively using *mean field theory*.

Let us approximate the interaction of each spin with its neighbors by an interaction with a mean-field, h:

$$H = -\sum_i h S_i^z \tag{3.73}$$

with h given by

$$h = \sum_i J \langle S_i^z \rangle = J z \langle S_i^z \rangle \tag{3.74}$$

where z is the coordination number. In this field, the partition function is just $2 \cosh \frac{h}{k_B T}$ and

$$\langle S^z \rangle = \tanh \frac{h}{k_B T} \tag{3.75}$$

Using the self-consistency condition, this is:

$$\langle S^z \rangle = \tanh \frac{J z \langle S^z \rangle}{k_B T} \tag{3.76}$$

For $k_B T < Jz$, this has non-zero solutions, $S^z \neq 0$ which break the symmetry $S_i^z \rightarrow -S_i^z$. In this *phase*, there is a spontaneous magnetization. For $k_B T > Jz$, there is only the solution $S^z = 0$. In this phase the symmetry is unbroken and the is no spontaneous magnetization. At $k_B T = Jz$, there is a critical point at which a phase transition occurs.

Chapter 4

Broken Translational Invariance in the Solid State

4.1 Simple Energetics of Solids

Why do solids form? The Hamiltonian of the electrons and ions is:

$$H = \sum_i \frac{p_i^2}{2m_e} + \sum_a \frac{P_a^2}{2M} + \sum_{i>j} \frac{e^2}{|r_i - r_j|} + \sum_{a>b} \frac{Z^2 e^2}{|R_a - R_b|} - \sum_{i,a} \frac{Ze^2}{|r_i - R_b|} \qquad (4.1)$$

It is invariant under the symmetry, $\vec{r}_i \to \vec{r}_i + \vec{a}$, $\vec{R}_a \to \vec{R}_a + \vec{a}$. However, the energy can usually be minimized by forming a crystal. At low enough temperature, this will win out over any possible entropy gain in a competing state, so crystallization will occur. Why is the crystalline state energetically favorable? This depends on the type of crystal. Different types of crystals minimize different terms in the Hamiltonian. In *molecular and ionic crystals*, the potential energy is minimized. In a molecular crystal, such as solid nitrogen, there is a van der Waals attraction between molecules caused by polarization of one by the other. The van der Waals attraction is balanced by short-range repulsion, thereby leading to a crystalline ground state. In an ionic crystal, such as NaCl, the electrostatic energy of the ions is minimized (one must be careful to take into account charge neutrality, without which the electrostatic

energy diverges). In *covalent and metallic crystals*, crystallization is driven by the minimization of the electronic kinetic energy. In a metal, such as sodium, the kinetic energy of the electrons is lowered by their ability to move throughout the metal. In a covalent solid, such as diamond, the same is true. The kinetic energy gain is high enough that such a bond can even occur between just two molecules (as in organic chemistry). The energetic gain of a solid is called the *cohesive energy*.

4.2   Phonons: Linear Chain

4.3   Quantum Mechanics of a Linear Chain

As a toy model of a solid, let us consider a linear chain of masses m connected by springs with spring constant B. Suppose that the equilibrium spacing between the masses is a. The equilibrium positions define a $1D$ lattice. The lattice 'vectors', $\vec{R}_j$, are defined by:

$$\vec{R}_j = ja \tag{4.2}$$

They connect the origin to all points of the lattice. If $\vec{R}$ and $\vec{R}'$ are lattice vectors, then $\vec{R} + \vec{R}'$ are also lattice vectors. A set of basis vectors is a minimal set of vectors which generate the full set of lattice vectors by taking linear combinations of the basis vectors. In our $1D$ lattice, a is the basis vector.

Let u_i be the displacement of the ith mass from its equilibrium position and let p_i be the corresponding momentum. Let us assume that there are N masses, and let's impose a periodic boundary condition, $u_i = u_{i+N}$. The Hamiltonian for such a system is:

$$H = \frac{1}{2m}\sum_i p_i^2 + \frac{1}{2}B\sum_i (u_i - u_{i+1})^2 \tag{4.3}$$

Let us use the Fourier transform representation:

$$
\begin{aligned}
u_j &= \frac{1}{\sqrt{N}} \sum_k u_k \, e^{ikja} \\
p_j &= \frac{1}{\sqrt{N}} \sum_k p_k \, e^{ikja}
\end{aligned}
\tag{4.4}
$$

As a result of the periodic boundary condition, the allowed k's are:

$$
k = \frac{2\pi n}{Na}
\tag{4.5}
$$

We can invert (4.4):

$$
\begin{aligned}
\frac{1}{\sqrt{N}} \sum_j u_j e^{ik'ja} &= \frac{1}{N} \sum_k \sum_j u_k \, e^{i(k-k')ja} \\
\frac{1}{\sqrt{N}} \sum_j u_j e^{ik'ja} &= u_{k'}
\end{aligned}
\tag{4.6}
$$

Note that $u_k^\dagger = u_{-k}$, $p_k^\dagger = p_{-k}$ since $u_j^\dagger = u_j$, $p_j^\dagger = p_j$. They satisfy the commutation relations:

$$
\begin{aligned}
[p_k, u_{k'}] &= \frac{1}{N} \sum_{j,j'} e^{ikja} e^{ik'j'a} \, [p_j, u_{j'}] \\
&= \frac{1}{N} \sum_{j,j'} e^{ikja} e^{ik'j'a} \, -i\hbar \delta_{jj'} \\
&= -i\hbar \frac{1}{N} \sum_j e^{i(k-k')ja} \\
&= -i\hbar \delta_{kk'}
\end{aligned}
\tag{4.7}
$$

Hence, p_k and $u_{k'}$ commute unless $k = k'$.

The displacements described by u_k are the same as those described by $u_{k+\frac{2\pi m}{a}}$ for any integer n:

$$
\begin{aligned}
u_{k+\frac{2\pi m}{a}} &= \frac{1}{\sqrt{N}} \sum_j u_j \, e^{-i\left(k+\frac{2\pi m}{a}\right)ja} \\
&= \frac{1}{\sqrt{N}} \sum_j u_j \, e^{-ikja} \\
&= u_k
\end{aligned}
\tag{4.8}
$$

Hence,

$$k \equiv k + \frac{2\pi m}{a} \tag{4.9}$$

Hence, we can restrict attention to $n = 0, 1, \ldots, N-1$ in (4.5). The Hamiltonian can be rewritten:

$$H = \sum_k \frac{1}{2m} p_k p_{-k} + 2B\left(\sin\frac{ka}{2}\right)^2 u_k u_{-k} \tag{4.10}$$

Recalling the solution of the harmonic oscillator problem, we define:

$$a_k = \frac{1}{\sqrt{2\hbar}}\left[\left(4mB\left(\sin\frac{ka}{2}\right)^2\right)^{\frac{1}{4}} u_k + \frac{i}{\left(4mB\left(\sin\frac{ka}{2}\right)^2\right)^{\frac{1}{4}}} p_k\right]$$

$$a_k^\dagger = \frac{1}{\sqrt{2\hbar}}\left[\left(4mB\left(\sin\frac{ka}{2}\right)^2\right)^{\frac{1}{4}} u_{-k} - \frac{i}{\left(4mB\left(\sin\frac{ka}{2}\right)^2\right)^{\frac{1}{4}}} p_{-k}\right] \tag{4.11}$$

(Recall that $u_k^\dagger = u_{-k}$, $p_k^\dagger = p_{-k}$.) which satisfy:

$$[a_k, a_k^\dagger] = 1 \tag{4.12}$$

Then:

$$H = \sum_k \hbar\omega_k \left(a_k^\dagger a_k + \frac{1}{2}\right) \tag{4.13}$$

with

$$\omega_k = 2\left(\frac{B}{m}\right)^{\frac{1}{2}} \left|\sin\frac{ka}{2}\right| \tag{4.14}$$

Hence, the linear chain is equivalent to a system of N independent harmonic oscillators. Its thermodynamics can be described by the Planck distribution.

The operators $a_k^\dagger, a_k$ are said to create and annihilate *phonons*. We say that a state $|\psi\rangle$ with

$$a_k^\dagger a_k |\psi\rangle = N_k |\psi\rangle \tag{4.15}$$

has N_k phonons of momentum k. Phonons are the quanta of lattice vibrations, analogous to photons, which are the quanta of oscillations of the electromagnetic field.

Observe that, as $k \to 0$, $\omega_k \to 0$:

$$
\begin{aligned}
\omega_{k \to 0} &= \left(\frac{Ba}{m/a}\right)^{\frac{1}{2}} k \\
&= \left(\frac{Ba}{\rho}\right)^{\frac{1}{2}} k
\end{aligned}
\tag{4.16}
$$

The physical reason for this is simple: an oscillation with $k = 0$ is a uniform transla-tion of the linear chain, which costs no energy.

Note that the reason for this is that the Hamiltonian is invariant under translations $u_i \to u_i + \lambda$. However, the ground state is not: the masses are located at the points $x_j = ja$. Translational invariance has been spontaneously broken. Of course, it could just as well be broken with $x_j = ja + \lambda$ and, for this reason, $\omega_k \to 0$ as $k \to 0$. An oscillatory mode of this type is called a *Goldstone mode*.

Consider now the case in which the masses are not equivalent. Let the masses alternate between m and M. As we will see, the *phonon spectra* will be more com-plicated. Let a be the distance between one m and the next m. The Hamiltonian is:

$$
H = \sum_i \left(\frac{1}{2m} p_{1,i}^2 + \frac{1}{2M} p_{2,i}^2 + \frac{1}{2} B \left(u_{1,i} - u_{2,i}\right)^2 + \frac{1}{2} B \left(u_{2,i} - u_{1,i+1}\right)^2\right)
\tag{4.17}
$$

The equations of motion are:

$$
\begin{aligned}
m \frac{d^2}{dt^2} u_{1,i} &= -B \left[(u_{1,i} - u_{2,i}) - (u_{2,i-1} - u_{1,i})\right] \\
M \frac{d^2}{dt^2} u_{2,i} &= -B \left[(u_{2,i} - u_{1,i}) + (u_{2,i} - u_{1,i+1})\right]
\end{aligned}
\tag{4.18}
$$

Going again to the Fourier representation, $\alpha = 1, 2$

$$
u_{\alpha,j} = \frac{1}{\sqrt{N}} \sum_k u_{\alpha,k} \, e^{ikja}
\tag{4.19}
$$

Where the allowed k's are:

$$
k = \frac{2\pi}{N} n
\tag{4.20}
$$

if there are $2N$ masses. As before,

$$u_{\alpha,k} = u_{\alpha,k+\frac{2\pi n}{a}} \tag{4.21}$$

$$\begin{pmatrix} m\frac{d^2}{dt^2} & 0 \\ 0 & M\frac{d^2}{dt^2} \end{pmatrix} \begin{pmatrix} u_{1,k} \\ u_{2,k} \end{pmatrix} = \begin{pmatrix} 2B & -B\left(1+e^{ika}\right) \\ -B\left(1+e^{-ika}\right) & 2B \end{pmatrix} \begin{pmatrix} u_{1,k} \\ u_{2,k} \end{pmatrix} \tag{4.22}$$

Fourier transforming in time:

$$\begin{pmatrix} -m\omega_k^2 & 0 \\ 0 & -M\omega_k^2 \end{pmatrix} \begin{pmatrix} u_{1,k} \\ u_{2,k} \end{pmatrix} = \begin{pmatrix} 2B & -B\left(1+e^{ika}\right) \\ -B\left(1+e^{-ika}\right) & 2B \end{pmatrix} \begin{pmatrix} u_{1,k} \\ u_{2,k} \end{pmatrix} \tag{4.23}$$

This eigenvalue equation has the solutions:

$$\omega_\pm^2 = B\left[\frac{1}{M} + \frac{1}{m} \pm \sqrt{\left(\frac{1}{M}+\frac{1}{m}\right)^2 - \frac{4}{nM}\left(\sin\frac{ka}{2}\right)^2} \right] \tag{4.24}$$

Observe that

$$\omega_{k\to 0}^- = \left(\frac{Ba}{2(m+M)/a}\right)^{\frac{1}{2}} k \tag{4.25}$$

This is the *acoustic* branch of the phonon spectrum in which m and M move in phase. As $k \to 0$, this is a translation, so $\omega_k^- \to 0$. Acoustic phonons are responsible for sound. Also note that

$$\omega_{k=\pi/a}^- = \left(\frac{2B}{M}\right)^{\frac{1}{2}} \tag{4.26}$$

Meanwhile, ω^+ is the *optical* branch of the spectrum (these phonons scatter light), in which m and M move in opposite directions.

$$\omega_{k\to 0}^+ = \sqrt{2B\left(\frac{1}{m}+\frac{1}{M}\right)} \tag{4.27}$$

$$\omega_{k=\pi/a}^+ = \left(\frac{2B}{m}\right)^{\frac{1}{2}} \tag{4.28}$$

so there is a gap in the spectrum of width

$$\omega_{\text{gap}} = \left(\frac{2B}{m}\right)^{\frac{1}{2}} - \left(\frac{2B}{M}\right)^{\frac{1}{2}} \tag{4.29}$$

Note that if we take $m = M$, we recover the previous results, with $a \to a/2$.

This is an example of what is called a *lattice with a basis*. Not every site on the chain is equivalent. We can think of the chain of $2N$ masses as a lattice with N sites. Each lattice site has a two-site basis: one of these sites has a mass m and the other has a mass M. Sodium Chloride is a simple subic lattice with a two-site basis: the sodium ions are at the vertices of an FCC lattice and the chlorine ions are displaced from them.

4.3.1   Statistical Mechnics of a Linear Chain

Let us return to the case of a linear chain of masses m separated by springs of force constant B, at equilibrium distance a. The excitations of this system are phonons which can have momenta $k \in [-\pi/a, \pi/a]$ (since $k \equiv k + \frac{2\pi m}{a}$), corresponding to energies

$$\hbar\omega_k = 2\left(\frac{B}{m}\right)^{\frac{1}{2}} \left|\sin\frac{ka}{2}\right| \tag{4.30}$$

Phonons are bosons whose number is not conserved, so they obey the Planck distribution. Hence, the energy of a linear chain at finite temperature is given by:

$$
\begin{aligned}
E &= \sum_k \frac{\hbar\omega_k}{e^{\beta\hbar\omega_k} - 1} \\
&= L \int_{-\frac{\pi}{a}}^{\frac{\pi}{a}} \frac{dk}{(2\pi)} \frac{\hbar\omega_k}{e^{\beta\hbar\omega_k} - 1}
\end{aligned}
\tag{4.31}
$$

Changing variables from k to ω,

$$
\begin{aligned}
E &= 2 \cdot \frac{L}{2\pi} \int_0^{\sqrt{4B/m}} \frac{2}{a} \frac{d\omega}{\sqrt{\frac{4B}{m} - \omega^2}} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \\
&= \frac{2N}{\pi} \int_0^{\sqrt{4B/m}} \frac{d\omega}{\sqrt{\frac{4B}{m} - \omega^2}} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \\
&= \frac{2N}{\pi} \frac{(k_B T)^2}{\hbar} \int_0^{\beta\hbar\sqrt{4B/m}} \frac{1}{\sqrt{\frac{4B}{m} - \left(\frac{x}{\beta\hbar}\right)^2}} \frac{x\, dx}{e^x - 1}
\end{aligned}
\tag{4.32}
$$

In the limit $k_B T \ll \hbar\sqrt{4B/m}$, we can take the upper limit of integration to ∞ and drop the x-dependent term in the square root in the denominator of the integrand:

$$E \quad = \quad \frac{N}{\pi}\sqrt{\frac{m}{4B}}\frac{(k_B T)^2}{\hbar}\int_0^\infty \frac{x\,dx}{e^x - 1}$$

$$(4.33)$$

Hence, $C_v \sim T$ at low temperatures.

In the limit $k_B T \gg \hbar\sqrt{4B/m}$, we can approximate $e^x \approx 1 + x$:

$$E \quad = \quad \frac{2N}{\pi}\frac{(k_B T)^2}{\hbar}\int_0^{\beta\hbar\sqrt{4B/m}} \frac{dx}{\sqrt{\frac{4B}{m} - \left(\frac{x}{\beta\hbar}\right)^2}}$$

$$= \quad N k_B \qquad\qquad (4.34)$$

In the intermediate temperature regime, a more careful analysis is necessary. In particular, note that the density of states, $1/\sqrt{\frac{4B}{m} - \omega^2}$ diverges at $\omega = \sqrt{4B/m}$; this is an example of a *van Hove singularity*. If we had alternating masses on springs, then the expression for the energy would have two integrals, one over the acoustic modes and one over the optical modes.

4.4   Lessons from the $1D$ chain

In the course of our analysis of the $1D$ chain, we developed the following strategy, which we will apply to crystals more generally in subsequent sections.

- Expand the Hamiltonian to Quadratic Order

- Fourier transform the Hamiltonian into momentum space

- Identify the Brillouin zone (range of distinct $\vec{k}$s)

- Rewrite the Hamiltonian in terms of creation and annihilation operators

- Obtain the Spectrum

- Compute the Density of States

- Use the Planck distribution to obtain the thermodynamics of the vibrational modes of the crystal.

4.5 Discrete Translational Invariance: the Reciprocal Lattice, Brillouin Zones, Crystal Momentum

Note that, in the above, momenta were only defined up to $\frac{2\pi n}{a}$. The momenta $\frac{2\pi n}{a}$ form a lattice in k-space, called the reciprocal lattice. This is true of any function which, like the ionic discplacements, is a function defined at the lattice sites. For such a function, $f\left(\vec{R}\right)$, defined on an arbitrary lattice, the Fourier transform

$$\tilde{f}\left(\vec{k}\right) = \sum_R e^{i\vec{k}\cdot\vec{R}} f\left(\vec{R}\right) \tag{4.35}$$

satisfies

$$\tilde{f}\left(\vec{k}\right) = \tilde{f}\left(\vec{k}+\vec{G}\right) \tag{4.36}$$

if $\vec{G}$ is in the set of reciprocal lattice vectors, defined by:

$$e^{i\vec{G}\cdot\vec{R}} = 1 \, , \ \ \text{for all } \vec{R} \tag{4.37}$$

The reciprocal lattice vectors also form a lattice since the sum of two reciprocal lattice vectors is also a reciprocal lattice vector. This lattice is called the *reciprocal lattice* or *dual lattice*.

In the analysis of the linear chain, we restricted momenta to $|k| < \pi/a$ to avoid double-counting of degrees of freedom. This restricted region in k-space is an example

of a *Brillouin zone* (or a first Brillouin zone). All of k-space can be obtained by translating the Brillouin zone through all possible reciprocal lattice vectors. We could have chosen our Brillouin zone differently by taking $0 < k < 2\pi/a$. Physically, there is no difference; the choice is a matter of convenience. What we need is a set of points in k space such that no two of these points are connected by a reciprocal lattice vector *and* such that all of k space can be obatained by translating the Brillouin zone through all possible reciprocal lattice vectors. We could even choose a Brillouin zone which is not connected, e.g. $0 < k < \pi/a$. or $3\pi/a < k < 4\pi/a$.

Later, we will consider solids with a more complicated lattice structure than our linear chain. Once again, phonon spectra will be defined in the Brillouin zone. Since $\tilde{f}\left(\vec{k}\right) = \tilde{f}\left(\vec{k} + \vec{G}\right)$, the phonon modes outside of the Brillouin zone are not physically distinct from those inside. One way of defining the Brillouin zone for an arbitrary lattice is to take all points in k space which are closer to the origin than to any other point of the reciprocal lattice. Such a choice of Brillouin zone is also called the Wigner-Seitz cell of the reciprocal lattice. We will discuss this in some detail later but, for now, let us consider the case of a cubic lattice. The lattice vectors of a cubic lattice of side a are:

$$\vec{R}_{n_1,n_2,n_3} = a\left(n_1\hat{x} + n_2\hat{y} + n_3\hat{z}\right) \tag{4.38}$$

The reciprocal lattice vectors are:

$$\vec{G}_{m_1,m_2,m_3} = \frac{2\pi}{a}\left(m_1\hat{x} + m_2\hat{y} + m_3\hat{z}\right) \tag{4.39}$$

The reciprocal lattice vectors also form a cubic lattice. The first Brillouin zone (Wigner-Seitz cell of the reciprocal lattice) is given by the cube of side $\frac{2\pi}{a}$ centered at the origin. The volume of this cube is related to the density according to:

$$\int_{\text{B.Z.}} \frac{d^3k}{(2\pi)^3} = \frac{1}{a^3} = \frac{N_{\text{ions}}}{V} \tag{4.40}$$

As we have noted before, the ground state (and the Hamiltonian) of a crystal is invariant under the discrete group of translations through all lattice vectors. Whereas

full translational invariance leads to momentum conservation, lattice translational symmetry leads to the conservation of *crystal momentum* – momentum up to a reciprocal lattice vector. (See Ashcroft and Mermin, appendix M.) For instance, in a collision between phonons, the difference between the incoming and outgoing phonon momenta can be any reciprocal lattice vector, G. Physically, one may think of the missing momentum as being taken by the lattice as a whole. This concept will also be important when we condsider the problem of electrons moving in the background of a lattice of ions.

4.6 Phonons: Continuum Elastic Theory

Consider the lattice of ions in a solid. Suppose the equilibrium positions of the ions are the sites $\vec{R}_i$. Let us describe small displacements from these sites by a displacement field $\vec{u}(\vec{R}_i)$. We will imagine that the crystal is just a system of masses connected by springs of equilibrium length a.

Before considering the details of the possible lattice structures of $2D$ and $3D$ crystals, let us consider the properties of a crystal at length scales which are much larger than the lattice spacing; this regime should be insensitive to many details of the lattice. At length scales much longer than its lattice spacing, a crystalline solid can be modelled as an elastic medium. We replace $\vec{u}(\vec{R}_i)$ by $\vec{u}(\vec{r})$ (i.e. we replace the lattice vectors, $\vec{R}_i$, by a continuous variable, $\vec{r}$). Such an approximation is valid at length scales much larger than the lattice spacing, a, or, equivalently, at wavevectors $q \ll 2\pi/a$.

In $1D$, we can take the continuum limit of our model of masses and springs:

$$
\begin{aligned}
H &= \frac{1}{2}m\sum_i\left(\frac{du_i}{dt}\right)^2 + \frac{1}{2}B\sum_i(u_i-u_{i+1})^2 \\
&= \frac{1}{2}\frac{m}{a}\sum_i a\left(\frac{du_i}{dt}\right)^2 + \frac{1}{2}Ba\sum_i a\left(\frac{u_i-u_{i+1}}{a}\right)^2
\end{aligned}
$$

$$\rightarrow \int dx \left(\frac{1}{2}\rho \left(\frac{du}{dt}\right)^2 + \frac{1}{2}\overline{B}\left(\frac{du}{dx}\right)^2 \right) \tag{4.41}$$

where ρ is the mass density and $\overline{B}$ is the bulk modulus. The equation of motion,

$$\frac{d^2 u}{dt^2} = \overline{B}\frac{d^2 u}{dx^2} \tag{4.42}$$

has solutions

$$u(x,t) = \sum_k u_k \, e^{i(kx-\omega t)} \tag{4.43}$$

with

$$\omega = \sqrt{\frac{\overline{B}}{\rho}}\, k \tag{4.44}$$

The generalization to a $3D$ continuum elastic medium is:

$$\rho\partial_t^2 \vec{u} = (\mu + \lambda)\,\vec{\nabla}\left(\vec{\nabla}\cdot\vec{u}\right) + \mu\nabla^2\vec{u} \tag{4.45}$$

where ρ is the mass density of the solid and μ and λ are the Lamé coefficients. Under a dilatation, $\vec{u}(\vec{r}) = \alpha\vec{r}$, the change in the energy density of the elastic medium is $\alpha^2(\lambda + 2\mu/3)/2$; under a shear stress, $u_x = \alpha y, u_y = u_z = 0$, it is $\alpha^2\mu/2$. In a crystal – which has only a discrete rotational symmetry – there may be more parameters than just μ and λ, depending on the symmetry of the lattice. In a crystal with cubic symmetry, for instance, there are, in general, three independent parameters. We will make life simple, however, and make the approximation of full rotational invariance.

The solutions are,

$$\vec{u}(\vec{r},t) = \vec{\epsilon}e^{i(\vec{k}\cdot\vec{r}-\omega t)} \tag{4.46}$$

where is $\vec{\epsilon}$ a unit polarization vector, satisfy

$$-\rho\omega^2\vec{\epsilon} = -(\mu + \lambda)\,\vec{k}\left(\vec{k}\cdot\vec{\epsilon}\right) - \mu k^2\vec{\epsilon} \tag{4.47}$$

For longitudinally polarized waves, $\vec{k} = k\vec{\epsilon}$,

$$\omega_k^l = \pm\sqrt{\frac{2\mu + \lambda}{\rho}}\, k \equiv \pm v_l k \tag{4.48}$$

while transverse waves, $\vec{k} \cdot \vec{\epsilon} = 0$ have

$$\omega_k^t = \pm \sqrt{\frac{\mu}{\rho}} \, k \equiv \pm v_s k \tag{4.49}$$

Above, we introduced the concept of the polarization of a phonon. In $3D$, the displacements of the ions can be in any direction. The two directions perpendicular to $\vec{k}$ are called *transverse*. Displacements along $\vec{k}$ are called *longitudinal*.

The Hamiltonian of this system,

$$H = \int d^3r \left(\frac{1}{2}\rho\big(\dot{\vec{u}}\big)^2 + \frac{1}{2}\left(\mu + \lambda\right) \left(\vec{\nabla} \cdot \vec{u}\right)^2 - \frac{1}{2}\mu\,\vec{u} \cdot \nabla^2 \vec{u}\right) \tag{4.50}$$

can be rewritten in terms of creation and annihilation operators,

$$\begin{aligned}
a_{k,s} &= \frac{1}{\sqrt{2\hbar}} \left[\sqrt{\rho\omega_{k,s}}\, \vec{\epsilon}_s \cdot \vec{u}_k + i\,\sqrt{\frac{\rho}{\omega_{k,s}}}\, \vec{\epsilon}_s \cdot \dot{\vec{u}}_k \right] \\
a_{k,s}^\dagger &= \frac{1}{\sqrt{2\hbar}} \left[\sqrt{\rho\omega_{k,s}}\, \vec{\epsilon}_s \cdot \vec{u}_{-k} - i\,\sqrt{\frac{\rho}{\omega_{k,s}}}\, \vec{\epsilon}_s \cdot \dot{\vec{u}}_{-k} \right]
\end{aligned} \tag{4.51}$$

as

$$H = \sum_{k,s} \hbar\omega_{k,s}\left(a_{k,s}^\dagger a_{k,s} + \frac{1}{2}\right) \tag{4.52}$$

Inverting the above definitions, we can express the displacement $\vec{u}(r)$ in terms of the creation and annihilation operators:

$$\vec{u}(\vec{r}) = \sum_{k,s} \sqrt{\frac{\hbar}{2\rho V \omega_k^s}}\, \vec{\epsilon}_s \left(a_{\vec{k},s} + a_{-\vec{k},s}^\dagger \right) e^{i\vec{k}\cdot\vec{r}} \tag{4.53}$$

$s = 1, 2, 3$ corresponds to the longitudinal and two transverse polarizations. Acting with $\vec{u}_k$ either annihilates a phonon of momentum k or creates a phonon of momentum $-k$.

The allowed $\vec{k}$ values are determined by the boundary conditions in a finite system. For periodic boundary conditions in a cubic system of size $V = L^3$, the allowed $\vec{k}$'s are $\frac{2\pi}{L}(n_1, n_2, n_3)$. Hence, the $\vec{k}$-space volume per allowed $\vec{k}$ is $(2\pi)^3/V$. Hence, we can take the infinite-volume limit by making the replacement:

$$\sum_k f(\vec{k}) = \frac{1}{(\Delta\vec{k})^3}\sum_k f(\vec{k})\,(\Delta\vec{k})^3$$

$$= \frac{V}{(2\pi)^3} \int d^3\vec{k} \, f(\vec{k}) \tag{4.54}$$

4.7 Debye theory

Since a solid can be modelled as a collection of independent oscillators, we can obtain the energy in thermal equilibrium using the Planck distribution function:

$$E = V \sum_s \int_{\text{B.Z.}} \frac{d^3 k}{(2\pi)^3} \frac{\hbar\omega_s(k)}{e^{\beta\hbar\omega_s(k)} - 1} \tag{4.55}$$

where $s = 1, 2, 3$ are the three polarizations of the phonons and the integral is over the Brillouin zone.

This can be rewritten in terms of the phonon density of states, $g(\omega)$ as:

$$E = V \int_0^\infty d\omega \, g(\omega) \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \tag{4.56}$$

where

$$g(\omega) = \sum_s \int_{\text{B.Z.}} \frac{d^3 k}{(2\pi)^3} \delta\left(\omega - \omega_s(k)\right) \tag{4.57}$$

The total number of states is given by:

$$
\begin{aligned}
\int_0^\infty d\omega \, g(\omega) &= \int_0^\infty d\omega \sum_s \int_{\text{B.Z.}} \frac{d^3 k}{(2\pi)^3} \delta\left(\omega - \omega_s(k)\right) \\
&= 3 \int_{\text{B.Z.}} \frac{d^3 k}{(2\pi)^3} \\
&= 3 \frac{N_{\text{ions}}}{V} \tag{4.58}
\end{aligned}
$$

The total number of normal modes is equal to the total number of ion degrees of freedom.

For a continuum elastic medium, there are two transverse modes with velocity v_t and one longitudinal mode with velocity v_l. In the limit that the lattice spacing is

very small, $a \to 0$, we expect this theory to be valid. In this limit, the Brillouin zone is all of momentum space, so

$$
\begin{aligned}
g_{\text{CEM}}(\omega) &= \int \frac{d^3 k}{(2\pi)^3} \left(2\delta \left(\omega - v_t k \right) \right) + \delta \left(\omega - v_l k \right))) \\
&= \frac{1}{2\pi^2} \left(\frac{2}{v_t^3} + \frac{1}{v_l^3} \right) \omega^2
\end{aligned}
\tag{4.59}
$$

In a crystalline solid, this will be a reasonable approximation to $g(\omega)$ for $k_B T \ll \hbar v_t / a$ where the only phonons present will be at low energies, far from the Brillouin zone boundary. At high temperatures, there will be thermally excited phonons near the Brillouin zone boundary, where the spectrum is definitely not linear, so we cannot use the continuum approximation. In particular, this $g(\omega)$ does not have a finite integral, which violates the condition that the integral should be the total number of degrees of freedom.

A simple approximation, due to Debye, is to replace the Brillouin zone by a sphere of radius k_D and assume that the spectrum is linear up to k_D. In other words, Debye assumed that:

$$
g_D(\omega) = \begin{cases} \frac{3}{2\pi^2 v^3} \omega^2 & \text{if } \omega < \omega_D \\ 0 & \text{if } \omega > \omega_D \end{cases}
$$

Here, we have assumed, for simplicity, that $v_l = v_t$ and we have written $\omega_D = v k_D$. ω_D is chosen so that

$$
\begin{aligned}
3 \frac{N_{\text{ions}}}{V} &= \int_0^\infty d\omega \, g(\omega) \\
&= \int_0^{\omega_D} d\omega \, \frac{3}{2\pi^2 v^3} \omega^2 \\
&= \frac{\omega_D^3}{2\pi^2 v^3}
\end{aligned}
\tag{4.60}
$$

i.e.

$$
\omega_D = \left(6\pi^2 v^3 N_{\text{ions}} / V \right)^{\frac{1}{3}}
\tag{4.61}
$$

With this choice,

$$
E = V \int_0^{\omega_D} d\omega \, \frac{3}{2\pi^2 v^3} \omega^2 \, \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}
$$

$$= V \frac{3(k_B T)^4}{2\pi^2 v^3 \hbar^3} \int_0^{\beta \hbar \omega_D} dx \, \frac{x^3}{e^x - 1} \tag{4.62}$$

In the low temperature limit, $\beta \hbar \omega_D \to \infty$, we can take the upper limit of the integral to ∞ and:

$$E \approx V \frac{3(k_B T)^4}{2\pi^2 v^3 \hbar^3} \int_0^{\infty} dx \, \frac{x^3}{e^x - 1} \tag{4.63}$$

The specific heat is:

$$C_V \approx T^3 \left[V \frac{12 k_B^4}{2\pi^2 v^3 \hbar^3} \int_0^{\infty} dx \, \frac{x^3}{e^x - 1} \right] \tag{4.64}$$

The T^3 contribution to the specific heat of a solid is often the most important contribution to the measured specific heat.

For $T \to \infty$,

$$
\begin{aligned}
E &\approx V \frac{3(k_B T)^4}{2\pi^2 v^3 \hbar^3} \int_0^{\beta \hbar \omega_D} dx \, x^2 \\
&= V \frac{\omega_D^3}{2\pi^2 v^3} k_B T \\
&= 3 N_{\text{ions}} k_B T
\end{aligned}
\tag{4.65}
$$

so

$$C_V \approx 3 N_{\text{ions}} k_B \tag{4.66}$$

The high-temperature specific heat is just $k_B/2$ times the number of degrees of freedom, as in classical statistical mechanics.

At high-temperature, we were guaranteed the right result since the density of states was normalized to give the correct total number of degrees of freedom. At low-temperature, we obtain a qualitatively correct result since the spectrum is linear. To obtain the exact result, we need to allow for longitudinal and transverse velocities which depend on the direction, $v_t\left(\hat{k}\right)$, $v_l\left(\hat{k}\right)$, since rotational invariance is not present. Debye's formula interpolates between these well-understood limits.

We can define θ_D by $k_B \theta_D = \hbar \omega_D$. For lead, which is soft, $\theta_D \approx 88K$, while for diamond, which is hard, $\theta_D \approx 1280K$.

4.8   More Realistic Phonon Spectra: Optical Phonons, van Hove Singularities

Although Debye's theory is reasonable, it clearly oversimplifies certain aspects of the physics. For instance, consider a crystal with a two-site basis. Half of the phonon modes will be optical modes. A crude approximation for the optical modes is an Einstein spectrum:

$$g_E(\omega) = \frac{N_{\text{ions}}}{2}\,\delta(\omega - \omega_E) \tag{4.67}$$

In such a case, the energy will be:

$$E = V\,\frac{3(k_B T)^4}{2\pi^2 v^3 \hbar^3}\int_0^{\beta\hbar\omega_{\text{max}}} dx\,\frac{x^3}{e^x - 1} + V\,\frac{N_{\text{ions}}}{2}\,\frac{\hbar\omega_E}{e^{\beta\hbar\omega_E} - 1} \tag{4.68}$$

with ω_{max} chosen so that

$$3\frac{N_{\text{ions}}}{2} = \frac{\omega_{\text{max}}^3}{2\pi^2 v^3} \tag{4.69}$$

Another feature missed by Debye's approximation is the existence of singularities in the phonon density of states. Consider the spectrum of the linear chain:

$$\omega(k) = 2\left(\frac{B}{m}\right)^{\frac{1}{2}}\left|\sin\frac{ka}{2}\right| \tag{4.70}$$

The minimum of this spectrum is at $k = 0$. Here, the density of states is well described by Debye theory which, for a $1D$ chain predicts $g(\omega) \sim$ const.. The maximum is at $k = \pi/a$. Near the maximum, Debye theory breaks down; the density of states is singular:

$$g(\omega) = \frac{2}{\pi a}\,\frac{1}{\sqrt{\omega_{\text{max}}^2 - \omega^2}} \tag{4.71}$$

In $3D$, the singularity will be milder, but still present. Consider a cubic lattice. The spectrum can be expanded about a maximum as:

$$\omega(k) = \omega_{\text{max}} - \alpha_x(k_x^{\text{max}} - k_x)^2 - \alpha_y\left(k_y^{\text{max}} - k_y\right)^2 - \alpha_z(k_z^{\text{max}} - k_z)^2 \tag{4.72}$$

Then (6 maxima; 1/2 of each ellipsoid is in the B.Z.)

$$
\begin{aligned}
G(\omega) &\equiv \int_\omega^{\omega_{\max}} d\omega\, g(\omega) \\
&= 6 \cdot \frac{1}{2} \cdot \frac{V}{(2\pi)^3} \left(\text{vol. of ellipsoid} \right) \\
&= 3 \frac{V}{(2\pi)^3} \frac{4}{3} \pi \frac{(\omega_{\max} - \omega)^{\frac{3}{2}}}{(\alpha_x \alpha_y \alpha_z)^{\frac{1}{2}}}
\end{aligned}
\tag{4.73}
$$

Differentiating:

$$
g(\omega) = \frac{3V}{4\pi^2} \frac{(\omega_{\max} - \omega)^{\frac{1}{2}}}{(\alpha_x \alpha_y \alpha_z)^{\frac{1}{2}}}
\tag{4.74}
$$

In $2D$ and $3D$, there can also be saddle points, where $\vec{\nabla}_k \omega(k) = 0$, but the eigenvalues of the second derivative matrix have different signs. At a saddle point, the phonon spectrum again has a square root singularity. van Hove proved that every $3D$ phonon spectrum has at least one maximum and two saddle points (one with one negative eigenvalue, one with two negative eigenvalues). To see why this might be true, draw the spectrum in the full k-space, repeating the Brillouin zone. Imagine drawing lines connecting the minima of the spectrum to the nearest neighboring minima (i.e. from each copy of the B.Z. to its neighbors). Imagine doing the same with the maxima. These lines intersect; at these intersections, we expect saddle points.

4.9 Lattice Structures

Thus far, we have focussed on general properties of the vibrational physics of crystalline solids. Real crystals come in a variety of different lattice structures, to which we now turn our attention.

4.9.1 Bravais Lattices

Bravais lattices are the underlying structure of a crystal. A $3D$ Bravais lattice is defined by the set of vectors $\vec{R}$

$$\left\{ \vec{R} \ \middle| \ \vec{R} = n_1 \vec{a}_1 + n_2 \vec{a}_2 + n_3 \vec{a}_3; \ n_i \in Z \right\} \tag{4.75}$$

where the vectors $\vec{a}_i$ are the *basis vectors* of the Bravais lattice. (Do not confuse with a lattice with a basis.) Every point of a Bravais lattice is equivalent to every other point. In an elemental crystal, it is possible that the elemental ions are located at the vertices of a Bravais lattice. In general, a crystal structure will be a Bravais lattice with a basis.

The symmetry group of a Bravais lattice is the group of translations through the lattice vectors together with some discrete rotation group about (any) one of the lattice points. In the problem set (Ashcroft and Mermin, problem 7.6) you will show that this rotation group can only have 2-fold, 3-fold, 4-fold, and 6-fold rotation axes.

There are 5 different types of Bravais lattice in $2D$: square, rectangular, hexagonal, oblique, and body-centered rectangular. There are 14 different types of Bravais lattices in $3D$. The $3D$ Bravais lattices are discussed in are described in Ashcroft and Mermin, chapter 7 (pp. 115-119). We will content ourselves with listing the Bravais lattices and discussing some important examples.

Bravais lattices can be grouped according to their symmetries. All but one can be obtained by deforming the cubic lattices to lower the symmetry.

- Cubic symmetry: cubic, FCC, BCC

- Tetragonal: stretched in one direction, $a \times a \times c$; tetragonal, centered tetragonal

- Orthorhombic: sides of 3 different lengths $a \times b \times c$, at right angles to each other; orthorhombic, base-centered, face-centered, body-centered.

- Monoclinic: One face is a parallelogram, the other two are rectangular; monoclinic, centered monoclinic.

- Triclinic: All faces are parallelograms.

- Trigonal: Each face is an $a \times a$ rhombus.

- Hexagonal: $2D$ hexagonal lattices of side a, stacked directly above one another, with spacing c.

Examples:

- Simple cubic lattice: $\vec{a}_i = a\,\hat{x}_i$.

- Body-centered cubic (BCC) lattice: points of a cubic lattice, together with the centers of the cubes $\cong$ interpenetrating cubic lattices offset by $1/2$ the body-diagonal.

$$\vec{a}_1 = a\,\hat{x}_1, \qquad \vec{a}_2 = a\,\hat{x}_2, \qquad \vec{a}_3 = \frac{a}{2}\,(\hat{x}_1 + \hat{x}_2 + \hat{x}_3) \tag{4.76}$$

Examples: Ba, Li, Na, Fe, K, Tl

- Face-centered cubic (FCC) lattice: points of a cubic lattice, together with the centers of the sides of the cubes, $\cong$ interpenetrating cubic lattices offset by $1/2$ a face-diagonal.

$$\vec{a}_1 = \frac{a}{2}\,(\hat{x}_2 + \hat{x}_3)\,, \qquad \vec{a}_2 = \frac{a}{2}\,(\hat{x}_1 + \hat{x}_3)\,, \qquad \vec{a}_3 = \frac{a}{2}\,(\hat{x}_1 + \hat{x}_2) \tag{4.77}$$

Examples: Al, Au, Cu, Pb, Pt, Ca, Ce, Ar.

- Hexagonal Lattice: Parallel planes of triangular lattices.

$$\vec{a}_1 = a\,\hat{x}_1, \qquad \vec{a}_2 = \frac{a}{2}\,\left(\hat{x}_1 + \sqrt{3}\,\hat{x}_2\right)\,, \qquad \vec{a}_3 = c\,\hat{x}_3 \tag{4.78}$$

Bravais lattices can be broken up into unit cells such that all of space can be recovered by translating a unit cell through all possible lattice vectors. A primitive unit cell is a unit cell of minimal volume. There are many possible choices of primitive unit cells. Given a basis, $\vec{a}_1, \vec{a}_2, \vec{a}_3$, a simple choice of unit cell is the region:

$$\left\{ \vec{r} \,\middle|\, \vec{r} = x_1\vec{a}_1 + x_2\vec{a}_2 + x_3\vec{a}_3; \ x_i \in [0,1] \right\} \tag{4.79}$$

The volume of this primitive unit cell and, thus, any primitive unit cell is:

$$\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3 \tag{4.80}$$

An alternate, symmetrical choice is the Wigner-Seitz cell: the set of all points which are closer to the origin than to any other point of the lattice. Examples: Wigner-Seitz for square=square, hexagonal=hexagon (not parallelogram), oblique=distorted hexagon, BCC=octohedron with each vertex cut off to give an extra square face (A+M p.74).

4.9.2   Reciprocal Lattices

If $\vec{a}_1, \vec{a}_2, \vec{a}_3$ span a Bravais lattice, then

$$
\begin{aligned}
\vec{b}_1 &= 2\pi \frac{\vec{a}_2 \times \vec{a}_3}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \\
\vec{b}_2 &= 2\pi \frac{\vec{a}_3 \times \vec{a}_1}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3} \\
\vec{b}_3 &= 2\pi \frac{\vec{a}_1 \times \vec{a}_2}{\vec{a}_1 \cdot \vec{a}_2 \times \vec{a}_3}
\end{aligned}
\tag{4.81}
$$

span the reciprocal lattice, which is also a bravais lattice.

The reciprocal of the reciprocal lattice is the set of all vectors $\vec{r}$ satisfying $e^{i\vec{G}\cdot\vec{r}} = 1$ for any recprocal lattice vector $\vec{G}$, i.e. it is the original lattice.

As we discussed above, a simple cubic lattice spanned by

$$a\hat{x}_1, \qquad a\hat{x}_2, \qquad a\hat{x}_3 \tag{4.82}$$

has the simple cubic reciprocal lattice spanned by:

$$\frac{2\pi}{a}\hat{x}_1 \,, \qquad \frac{2\pi}{a}\hat{x}_2 \,, \qquad \frac{2\pi}{a}\hat{x}_3 \tag{4.83}$$

An FCC lattice spanned by:

$$\frac{a}{2}\left(\hat{x}_2 + \hat{x}_3\right) \,, \qquad \frac{a}{2}\left(\hat{x}_1 + \hat{x}_3\right) \,, \qquad \frac{a}{2}\left(\hat{x}_1 + \hat{x}_2\right) \tag{4.84}$$

has a BCC reciprocal lattice spanned by:

$$\frac{4\pi}{a}\frac{1}{2}\left(\hat{x}_2 + \hat{x}_3 - \hat{x}_1\right) \,, \qquad \frac{4\pi}{a}\frac{1}{2}\left(\hat{x}_1 + \hat{x}_3 - \hat{x}_2\right) \,, \qquad \frac{4\pi}{a}\frac{1}{2}\left(\hat{x}_1 + \hat{x}_2 - \hat{x}_3\right) \tag{4.85}$$

Conversely, a BCC lattice has an FCC reciprocal lattice.

The Wigner-Seitz primitive unit cell of the reciprocal lattice is the *first Brillouin zone.* In the problem set (Ashcroft and Mermin, problem 5.1), you will show that the Brillouin zone has volume $(2\pi)^3/v$ if the volume of the unit cell of the original lattice is v. The first Brillouin zone is enclosed in the planes which are the perpendicular bisectors of the reciprocal lattice vectors. These planes are called *Bragg planes* for reasons which will become clear below.

4.9.3 Bravais Lattices with a Basis

Most crystalline solids are not Bravais lattices: not every ionic site is equivalent to every other. In a compound this is necessarily true; even in elemental crystals it is often the case that there are inequivalent sites in the crystal structure. These crystal structures are *lattices with a basis.* The classification of such structures is discused in Ashcroft and Mermin, chapter 7 (pp. 119-126). Again, we will content ourselves with discussing some important examples.

- Honeycomb Lattice (2D): A triangular lattice with a two-site basis. The triangular lattice is spanned by:

$$\vec{a}_1 = \frac{a}{2}\left(\sqrt{3}\,\hat{x}_1 + 3\,\hat{x}_2\right) \,, \qquad \vec{a}_2 = a\sqrt{3}\,\hat{x}_1 \tag{4.86}$$

The two-site basis is:

$$\mathbf{0} , \qquad a\,\hat{x}_2 \tag{4.87}$$

Example: Graphite

- Diamond Lattice: FCC lattice with a two-site basis: The two-site basis is:

$$\mathbf{0} , \qquad \frac{a}{4}\left(\hat{x}_1 + \hat{x}_2 + \hat{x}_3\right) \tag{4.88}$$

Example: Diamond, Si, Ge

- Hexagonal Close-Packed (HCP): Hexagonal lattice with a two-site basis:

$$\mathbf{0} , \qquad \frac{a}{2}\,\hat{x}_1 + \frac{a}{2\sqrt{3}}\,\hat{x}_2 + \frac{c}{2}\,\hat{x}_3 \tag{4.89}$$

Examples: Be,Mg,Zn, ...

- Sodium Chloride: Cubic lattice with Na and Cl at alternate sites $\cong$ FCC lattice with a two-site basis:

$$\mathbf{0} , \qquad \frac{a}{2}\left(\hat{x}_1 + \hat{x}_2 + \hat{x}_3\right) \tag{4.90}$$

Examples: NaCl,NaF,KCl

4.10 Bragg Scattering

One way of experimentally probing a condensed matter system involves scattering a photon or neutron off the system and studying the energy and angular dependence of the resulting cross-section. Crystal structure experiments have usually been done with X-rays.

Let us first examine this problem intuitively and then in a more systematic fashion. Consider, first, elastic scattering of X-rays. Think of the X-rays as photons which can take different paths through the crystal. Consider the case in which $\vec{k}$ is the

wavevector of an incoming photon and $\vec{k}'$ is the wavevector of an outgoing photon. Let us, furthermore, assume that the photon only scatters off one of the atoms in the crystal (the probability of multiple scattering is very low). This atom can be any one of the atoms in the crystal. These different scattring events will interfere constructively if the path lengths differ by an integer number of wavelengths. The extra path length for a scattering off an atom at $\vec{R}$, as compared to an atom at the origin is:

$$\vec{R} \cdot \vec{k} + \left(-\vec{R} \cdot \vec{k}'\right) = \vec{R} \cdot \left(\vec{k} - \vec{k}'\right) \tag{4.91}$$

If this is an integral multiple of 2π for all lattice vectors $\vec{R}$, then scattering interferes constructively. By definition, this implies that $\vec{k} - \vec{k}'$ must be a reciprocal lattice vector. For elastic scatering, $|\vec{k}| = |\vec{k}'|$, so this implies that there is a reciprocal lattic vector $\vec{G}$ of magnitude

$$\left|\vec{G}\right| = 2|\vec{k}| \sin \theta \tag{4.92}$$

where θ is the angle between the incoming and outgoing X-rays.

To rederive this result more formally, let us assume that our crystal is in thermal equilibrium at inverse temperature β and that photons interact with our crystal via the Hamiltonian H'. Suppose that photons of momentum $\vec{k}_i$, and energy ω_i are scattered by our system. The differential cross-section for the photons to be scattered into a solid angle $d\Omega$ centered about $\vec{k}_f$ and into the energy range between $\omega_f \pm d\omega$ is:

$$\frac{d^2\sigma}{d\Omega \, d\omega} = \sum_{m,n} \frac{k_f}{k_i} \left|\left\langle \vec{k}_f; m \left| H' \right| \vec{k}_i; n \right\rangle\right|^2 e^{-\beta E_n} \, \delta\left(\omega + E_n - E_m\right) \tag{4.93}$$

where $\omega = \omega_i - \omega_f$ and n and m label the initial and final states of our crystal. Let $\vec{q} = \vec{k}_f - \vec{k}_i$.

Let us assume that the interactions between the photon and the ions in our system is of the form:

$$H' = \sum_R U\left(\vec{x} - \left(\vec{R} + \vec{u}\left(\vec{R}\right)\right)\right) \tag{4.94}$$

Then

$$
\begin{aligned}
\left\langle \vec{k}_f; m \,|H'|\, \vec{k}_i; n \right\rangle
&= \int d^3x \, \frac{1}{V} e^{i\vec{q}\cdot\vec{x}} \left\langle m \left| \sum_R U\left(\vec{x} - \left(\vec{R} + \vec{u}\left(\vec{R}\right)\right)\right) \right| n \right\rangle \\
&= \left\langle m \left| \frac{1}{V} \sum_R e^{-i\vec{q}\cdot\left(\vec{R}+\vec{u}(\vec{R})\right)} \right| n \right\rangle \tilde{U}(\vec{q}) \\
&= \frac{1}{V} \sum_R \left[e^{-i\vec{q}\cdot\vec{R}} \left\langle m \left| e^{-i\vec{q}\cdot\vec{u}(\vec{R})} \right| n \right\rangle \right] \tilde{U}(\vec{q}) \\
&= \left[\frac{1}{V} \sum_R e^{-i\vec{q}\cdot\vec{R}} \right] \left\langle m \left| e^{-i\vec{q}\cdot\vec{u}(0)} \right| n \right\rangle \tilde{U}(\vec{q})
\end{aligned}
\tag{4.95}
$$

Let us consider, first, the case of elastic scattering, in which the state of the crystal does not change. Then $|n\rangle = |m\rangle$, $|k_i| = |k_f| \equiv k$, $|q| = 2k \sin\frac{\theta}{2}$, and:

$$
\left\langle \vec{k}_f; n \,|H'|\, \vec{k}_i; n \right\rangle = \left[\frac{1}{V} \sum_R e^{-i\vec{q}\cdot\vec{R}} \right] \left\langle n \left| e^{-i\vec{q}\cdot\vec{u}(0)} \right| n \right\rangle \tilde{U}(\vec{q})
\tag{4.96}
$$

Let us focus on the sum over the lattice:

$$
\frac{1}{V} \sum_R e^{-i\vec{q}\cdot\vec{R}} = \sum_{\text{R.L.V. } G} \delta_{\vec{q},\vec{G}}
\tag{4.97}
$$

The sum is 1 if $\vec{q}$ is a reciprocal lattice vector and vanishes otherwise. The scattering cross-section is given by:

$$
\frac{d^2\sigma}{d\Omega d\omega} = \sum_n e^{-\beta E_n} \left| \left\langle n \left| e^{-i\vec{q}\cdot\vec{u}(0)} \right| n \right\rangle \right|^2 \left| \tilde{U}(\vec{q}) \right|^2 \left[\sum_{\text{R.L.V. } G} \delta_{\vec{q},\vec{G}} \right]
\tag{4.98}
$$

In other words, the scattering cross-section is peaked when the photon is scattered through a reciprocal lattice vector $\vec{k}_f = \vec{k}_i + \vec{G}$. For elastic scattering, this requires

$$
\left(\vec{k}_i\right)^2 = \left(\vec{k}_i + \vec{G}\right)^2
\tag{4.99}
$$

or,

$$
\left(\vec{G}\right)^2 = -2\,\vec{k}_i \cdot \vec{G}
\tag{4.100}
$$

This is called the *Bragg condition*. It is satisfied when the endpoint of $\vec{k}$ is on a Bragg plane. When it is satisfied, *Bragg scattering* occurs.

When there is structure within the unit cell, as in a lattice with a basis, the formula is slightly more complicated. We can replace the photon-ion interaction by:

$$H' = \sum_R \sum_b U_b \left(\vec{x} - \left(\vec{R} + \vec{b} + \vec{u}(\vec{R} + \vec{b}) \right) \right) \tag{4.101}$$

Then,

$$\frac{1}{V} \sum_R e^{-i\vec{q}\cdot\vec{R}} \, \tilde{U}(\vec{q}) \tag{4.102}$$

is replaced by

$$\frac{1}{V} \sum_R f_q \, e^{-i\vec{q}\cdot\vec{R}} \tag{4.103}$$

where

$$f_q = \sum_b U_b \left(\vec{x} \right) e^{-i\vec{q}\cdot\vec{x}} \tag{4.104}$$

As a result of the *structure factor*, f_q, the scattering amplitude need not have a peak at every reciprocal lattic vector, $\vec{q}$.

Of course, the probability that the detector is set up at precisely the right angle to receive $\vec{k}_f = \vec{k}_i + \vec{G}$ is very low. Hence, these experiments are usually done with a powder so that there will be Bragg scattering whenever $2k \sin \frac{\theta}{2} = |G|$. By varying θ, a series of peaks are seen at, e.g. $\pi/6$, $\pi/4$, etc., from which the reciprocal lattice vectors are reconstructed.

Since $|k| \sim |G| \sim \left(1\text{Å} \right)^{-1}$, the energy of the incoming photons is $\sim \hbar ck \sim 10^4 eV$ which is definitely in the X-ray range.

Thus far, we have not looked closely at the factor:

$$\left| \left\langle m \left| e^{-i\vec{q}\cdot\vec{u}(0)} \right| n \right\rangle \right|^2 \tag{4.105}$$

This factor results from the vibration of the lattice due to phonons. In elastic scattering, the amplitude of the peak will be reduced by this factor since the probability of the ions forming a perfect lattice is less than 1. The inelastic amplitude will contain contributions from processes in which the incoming photon or neutron creates a

phonon, thereby losing some energy. By measuring inelastic neutron scattering (for which the energy resolution is better than for X-rays), we can learn a great deal about the phonon spectrum.

Chapter 5

Electronic Bands

5.1 Introduction

Thus far, we have ignored the dynamics of the elctrons and focussed on the ionic vibrations. However, the electrons are important for many properties of solids. In metals, the specific heat is actually $C_V = \gamma T + \alpha T^3$. The linear term is due to the electrons. Electrical conduction is almost always due to the electrons, so we will need to understand the dynamics of electrons in solids in order to compute, for instance, the conductivity $\sigma(T, \omega)$.

In order to do this, we will need to understand the quantum mechanics of electrons in the periodic potential due to the ions. Such an analysis will enable us to understand some broad features of the electronic properties of crystalline solids, such as the distinction between metals and insulators.

5.2 Independent Electrons in a Periodic Potential: Bloch's theorem

Let us first neglect all interactions between the electrons and focus on the interactions between each electron and the ions. This may seem crazy since the inter-electron

interaction isn't small, but let us make this approximation and proceed. At some level, we can say that we will include the electronic contribution to the potential in some average sense so that the electrons move in the potential created by the ions and by the average electron density (of course, we shoudld actually do this self-consistently). Later, we will see why this is sensible.

When the electrons do not interact with each other, the many-electron wavefunction can be constructed as a Slater determinant of single-electron wavefunctions.

Hence, we have reduced the problem to that of a single electron moving in a lattice of ions. The Hamiltonian for such a problem is:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + \sum_R V(\vec{x} - \vec{R} - \vec{u}(\vec{R})) \tag{5.1}$$

expanding in powers of $\vec{u}(\vec{R})$,

$$H = -\frac{\hbar^2}{2m}\nabla^2 + \sum_R V(\vec{x} - \vec{R}) - \sum_R \vec{\nabla}V(\vec{x} - \vec{R}) \cdot \vec{u}(\vec{R}) + \ldots \tag{5.2}$$

The third term and the . . . are electron-phonon interaction terms. They can be treated as a perturbation. We will focus on the first two terms, which describe an electron moving in a periodic potential. This highly simplified problem already contains much of the qualitative physics of a solid.

Let us begin by proving an important theorem due to Bloch.

Bloch's Theorem: If $V(\vec{r} + \vec{R}) = V(\vec{r})$ for all lattice vectors $\vec{R}$ of some given lattice, then for any solution of the Schrödinger equation in this potential,

$$-\frac{\hbar^2}{2m}\nabla^2\psi(\vec{r}) + V(\vec{r})\psi(\vec{r}) = E\psi(\vec{r}) \tag{5.3}$$

there exists a $\vec{k}$ such that

$$\psi(\vec{r} + \vec{R}) = e^{i\vec{k}\cdot\vec{R}}\psi(\vec{r}) \tag{5.4}$$

Proof: Consider the lattice translation operator T_R which acts according to

$$T_R\chi(\vec{r}) = \chi(\vec{r} + \vec{R}) \tag{5.5}$$

Then

$$T_R H \chi(\vec{r}) = H T_R \chi(\vec{r} + \vec{R}) \tag{5.6}$$

i.e. $[T_R, H] = 0$. Hence, we can take our energy eigenstates to be eigenstates of T_R. Hence, for any energy eigenstate $\psi(\vec{r})$,

$$T_R \psi(\vec{r}) = c(\vec{R}) \psi(\vec{r}) \tag{5.7}$$

The additivity of the translation group implies that,

$$c(\vec{R}) c(\vec{R'}) = c(\vec{R} + \vec{R'}) \tag{5.8}$$

Hence, there is some k such that

$$c(\vec{R}) = e^{i\vec{k}\cdot\vec{R}} \tag{5.9}$$

Since $e^{i\vec{G}\cdot\vec{R}} = 1$ if $\vec{G}$ is a reciprocal lattice vector, we can always take $\vec{k}$ to be in the first Brillouin zone.

5.3   Tight-Binding Models

Let's consider a very simple model of a $1D$ solid in which we imagine that the atomic nuclei lie along a chain of spacing a. Consider a single ion and focus on two of its electronic energy levels. In real systems, we will probably consider s and d orbitals, but this is not important here; in our toy model, these are simply two electronic states which are localized about the atomic nucleus. We'll call them $|1\rangle$ and $|2\rangle$, with energies ϵ_1^0 and ϵ_2^0. Let's further imagine that the splitting $\epsilon_2^0 - \epsilon_1^0$ between these levels is large. Now, when we put this atom in the linear chain, there will be some overlap between these levels and the corresponding energy levels on neighboring atoms. We can model such a system by the Hamiltonian:

$$H = \sum_R \left(\epsilon_0^1 |R,1\rangle\langle R,1| + \epsilon_0^2 |R,2\rangle\langle R,2| \right) - \sum_{R,R' \text{ n.n}} (t_1 |R,1\rangle\langle R',1| + t_2 |R,2\rangle\langle R',2|) \tag{5.10}$$

We have assumed that t_1 is the amplitude for an electron at $|R, 1\rangle$ to hop to $|R', 1\rangle$, and similarly for t_2. For simplicity, we have ignored the possibility of hopping from $|R, 1\rangle$ to $|R', 2\rangle$, which is unimportant anyway when ϵ_0 is large. The eigenstates of this Hamiltonian are:

$$|k, 1\rangle = \sum_R e^{ikR}|R, 1\rangle \tag{5.11}$$

with energy

$$\epsilon_1(k) = \epsilon_1^0 - 2t_1 \cos ka \tag{5.12}$$

and

$$|k, 2\rangle = \sum_R e^{ikR}|R, 2\rangle \tag{5.13}$$

with energy

$$\epsilon_2(k) = \epsilon_2^0 - 2t_2 \cos ka \tag{5.14}$$

Note, first, that k lives in the first Brillouin zone since

$$|k, i\rangle \equiv \left|k + \frac{2\pi n}{a}, i\right\rangle \tag{5.15}$$

Now, observe that the two atomic energy levels have broadened into two *energy bands*. There is a *band gap* between these bands of magnitude $\epsilon_2^0 - \epsilon_1^0 - 2t_1 - 2t_2$. This is a characteristic feature of electronic states in a periodic potential: the states break up into *bands* with *energy gaps* separating the bands.

How many states are there in each band? As we discussed in the context of phonons, there are as many allowed k's in the Brillouin zone as there are ions in the crystal. Let's repeat the argument. The Brillouin zone has k-space extent $2\pi/a$. In a finite-size system of length L with periodic boundary conditions, allowed k's are of the form $2\pi n/L$ where n is an integer. Hence, there are $L/a = N_{\text{ions}}$ allowed k's in the Brillouin zone. (This argument generalizes to arbitrary lattices in arbitrary dimension.) Hence, there are as many states as lattice sites. Each state can be filled by one up-spin electon and one down-spin electron. Hence, if the atom is monovalent

– i.e. if there is one electron per site – then, in the ground state, the lower band, $|k, 1\rangle$ is half-filled and the upper band is empty. The Fermi energy is at ϵ_1^0. The Fermi momentum (or, more properly, Fermi crystal momentum) is at $\pm \pi/2a$. At low temperature, the fact that there is a gap far away from the Fermi momentum is unimportant, and the Fermi sea will behave just like the Fermi sea of a free Fermi gas. In particular, there is no energy gap in the many-electron spectrum since we can always excite an electron from a filled state just below the Fermi surface to one of the unfilled states just above the gap. For instance, the electronic contribution to the specific heat will be $C_V \sim T$. The difference is that the density of states will be different from that of a free Fermi gas. In situations such as this, when a band is partially filled, the crystal is (almost always) a metal. (Sometimes inter-electron interactions can make such a system an insulator.)

If there are two electrons per lattice site, then the lower band is filled and the upper band is empty in the ground state. In such a case, there is an energy gap $E_g = \epsilon_2^0 - \epsilon_1^0 - 2t_1 - 2t_2$ between the ground state and the lowest excited state which necessarily involves exciting an electron from the lower band to the upper band. Crystals of this type, which have no partially filled bands, are insulators. The electronic contribution to the specific heat will be suppressed by a factor of $e^{-E_g/T}$.

Note that the above *tight-binding model* can be generalized to arbitrary dimension of lattice. For instance, a cubic lattice with one orbital per site has tight-binding spectrum:

$$\epsilon(k) = -2t \left(\cos k_x a + \cos k_y a + \cos k_z a\right) \tag{5.16}$$

Again, if there is one electron per site, the band will be half-filled (and metallic); if there are two electrons per site the band will be filled (and insulating).

The model which we have just examined is grossly oversimplified but can, nevertheless, be justified, to an extent. Let us reconsider our lattice of atoms.

Electronic orbitals of an isolated atom:

$$\varphi_n(\vec{r}) \tag{5.17}$$

with energies ϵ_n:

$$-\frac{\hbar^2}{2m}\nabla^2\varphi_n(\vec{r}) + V(\vec{r})\varphi_n(\vec{r}) = \epsilon_n\varphi_n(\vec{r}) \tag{5.18}$$

We now want to solve:

$$-\frac{\hbar^2}{2m}\nabla^2\psi_k(\vec{r}) + \sum_R V(\vec{r}+\vec{R})\,\psi_k(\vec{r}) = \epsilon(k)\psi_k(\vec{r}) \tag{5.19}$$

Let's try the ansatz:

$$\psi_k(\vec{r}) = \sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\,\varphi_n(\vec{r}+\vec{R}) \tag{5.20}$$

which satisfies Bloch's theorem. Substituting into Schrödinger's equation and taking the matrix element with φ_m, we get:

$$\int d^3r\, \varphi_m^*(\vec{r}) \left(-\tfrac{\hbar^2}{2m}\nabla^2 + \sum_{R'} V(\vec{r}+\vec{R}')\right) \sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\,\varphi_n(\vec{r}+\vec{R}) =$$
$$\epsilon(k)\int d^3r\, \varphi_m^*(\vec{r}) \sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\,\varphi_n(\vec{r}+\vec{R}) \tag{5.21}$$

Let's write

$$-\frac{\hbar^2}{2m}\nabla^2 + \sum_{R'} V(\vec{r}+\vec{R}') = -\frac{\hbar^2}{2m}\nabla^2 + V(\vec{r}+\vec{R}) + \sum_{R'\neq R} V(\vec{r}+\vec{R}')$$
$$= H_{\text{at},R} + \Delta V_R(r) \tag{5.22}$$

Then, we have

$$\int d^3r\, \varphi_m^*(\vec{r}) \sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\left(H_{\text{at},R} + \Delta V_R(r)\right)\varphi_n(\vec{r}+\vec{R}) = \epsilon(k)\int d^3r\, \varphi_m^*(\vec{r})\sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\,\varphi_n(\vec{r}+\vec{R}) \tag{5.23}$$

$$\sum_{R,n} c_n\, \epsilon_n e^{i\vec{k}\cdot\vec{R}}\int d^3r\, \varphi_m^*(\vec{r})\varphi_n(\vec{r}+\vec{R}) + \sum_{R,n} c_n\, e^{i\vec{k}\cdot\vec{R}}\int d^3r\, \varphi_m^*(\vec{r})\Delta V_R(r)\varphi_n(\vec{r}+\vec{R}) =$$
$$\epsilon(k)c_m + \epsilon(k)\sum_{R\neq 0,n} e^{i\vec{k}\cdot\vec{R}} c_n \int d^3r\, \varphi_m^*(\vec{r})\varphi_n(\vec{r}+\vec{R}) \tag{5.24}$$

$$c_m \epsilon_m + \sum_{R \neq 0,n} c_n \epsilon_n e^{i\vec{k}\cdot\vec{R}} \int d^3r \, \varphi_m^*(\vec{r})\varphi_n(\vec{r}+\vec{R}) \; +$$

$$\sum_{R,n} c_n e^{i\vec{k}\cdot\vec{R}} \int d^3r \, \varphi_m^*(\vec{r}) \left[\Delta V_R(r)\right] \varphi_n(\vec{r}+\vec{R}) =$$

$$\epsilon(k)c_m + \epsilon(k) \sum_{R \neq 0,n} e^{i\vec{k}\cdot\vec{R}} c_n \int d^3r \, \varphi_m^*(\vec{r})\varphi_n(\vec{r}+\vec{R}) \quad (5.25)$$

Writing:

$$\alpha_{mn}(\vec{R}) = \int d^3r \, \varphi_m^*(\vec{r})\varphi_n(\vec{r}+\vec{R})$$

$$\gamma_{mn}(\vec{R}) = -\int d^3r \, \varphi_m^*(\vec{r}) \left[\Delta V_R(r)\right] \varphi_n(\vec{r}+\vec{R}) \quad (5.26)$$

We have:

$$c_m \left(\epsilon_m - \epsilon(k)\right) + \sum_{R \neq 0,n} c_n \left(\epsilon_n - \epsilon(k)\right) e^{i\vec{k}\cdot\vec{R}} \alpha_{mn}(\vec{R}) = \sum_{R,n} c_n e^{i\vec{k}\cdot\vec{R}} \gamma_{mn}(\vec{R}) \quad (5.27)$$

Both $\alpha_{mn}(\vec{R})$ and $\gamma_{mn}(\vec{R})$ are exponentially small, $\sim e^{-R/a_0}$. In particular, $\alpha_{mn}(\vec{R})$ and $\gamma_{mn}(\vec{R})$ are much larger for nearest neighbors than for any other sites, so let's neglect the other matrix elements and write $\alpha_{mn} = \alpha_{mn}(\vec{R}_{\text{n.n.}})$, $\gamma_{mn} = \gamma_{mn}(\vec{R}_{\text{n.n.}})$, $v_{mn} = \gamma_{mn}(0)$. In problem 2 of problem set 7, so may make these approximations. Suppose that we make the approximation that the lth orbital is well separated in energy from the others. Then we can neglect $\alpha_{ln}(\vec{R})$ and $\gamma_{ln}(\vec{R})$ for $n \neq l$. We write $\beta = v_{ll}$. Focusing on the $m = l$ equation, we have:

$$\left(\epsilon_l - \epsilon(k)\right) + \alpha_{ll} \left(\epsilon_l - \epsilon(k)\right) \sum_{R\text{n.n.}} e^{i\vec{k}\cdot\vec{R}} = \beta + \gamma_{ll} \sum_{R\text{n.n.}} e^{i\vec{k}\cdot\vec{R}} \quad (5.28)$$

Hence,

$$\epsilon(k) = \epsilon_l - \frac{\beta + \gamma_{ll} \sum_{R\text{n.n.}} e^{i\vec{k}\cdot\vec{R}}}{1 + \alpha_{ll} \sum_{R\text{n.n.}} e^{i\vec{k}\cdot\vec{R}}} \quad (5.29)$$

If we neglect the α's and retain only the γ's, then we recover the result of our phenomenological model. For instance, for the cubic lattice, we have:

$$\epsilon(k) = [\epsilon_l - \beta] - 2\gamma_{ll} \left[\cos k_x a + \cos k_y a + \cos k_z a\right] \quad (5.30)$$

Tight-binding models give electronic wavefunctions which are a coherent super-position of localized atomic orbitals. Such wavefunctions have very small amplitude

in the interstitial regions between the ions. Such models are valid, as we have seen, when there is very little overlap between atomic wavefunctions on neighboring atoms. In other words, a tight-binding model will be valid when the size of an atomic orbital is smaller than the interatomic distance, i.e. $a_0 \ll R$. In the case of core electrons, e.g. $1s$, $2s$, $2p$, this is the case. However, this is often not the case for valence electrons, e.g. $3s$ electrons. Nevertheless, the tight-binding method is a simple method which gives many qualitative features of electronic bands. In the study of high-T_c superconductivity, it has proven useful for this reason.

5.4 The δ-function Array

Let us now consider another simple toy-model of a solid, a $1D$ array of δ-functions:

$$\left(-\frac{\hbar^2}{2m}\frac{d^2}{dx^2} + V \sum_{n=-\infty}^{\infty} \delta(x-na) \right) \psi(x) = E\,\psi(x) \tag{5.31}$$

Between the peaks of the δ functions, $\psi(x)$ must be a superposition of the plane waves e^{iqx} and e^{-iqx} with energy $E(q) = \hbar^2 q^2/2m$. Between $x = 0$ and $x = a$,

$$\psi(x) = e^{iqx+i\alpha} + e^{-iqx-i\alpha} \tag{5.32}$$

with α complex. According to Bloch's theorem,

$$\psi(x+a) = e^{ika}\,\psi(x) \tag{5.33}$$

Hence, in the region between $x = a$ and $x = 2a$,

$$\psi(x) = e^{ika}\left(e^{iq(x-a)+i\alpha} + e^{-iq(x-a)-i\alpha} \right) \tag{5.34}$$

Note that k which determines the transformation property under a translation $x \to x + a$ is *not* the same as q, which is the 'local' momentum of the electron, which determines the energy. Continuity at $x = a$ implies that

$$\cos(qa + \alpha) = e^{ika}\cos\alpha \tag{5.35}$$

or,

$$\tan \alpha = \frac{\cos qa - e^{ika}}{\sin qa} \tag{5.36}$$

Integrating Schrödinger's equation from $x = a - \epsilon$ to $x = a + \epsilon$, we have,

$$\sin(qa + \alpha) - e^{ika} \sin \alpha = \frac{2mV}{\hbar^2 q} e^{ika} \cos \alpha \tag{5.37}$$

or,

$$\tan \alpha = \frac{\frac{2mV}{\hbar^2 q} e^{ika} - \sin qa}{\cos qa - e^{ika}} \tag{5.38}$$

Combining these equations,

$$e^{2ika} - 2 \left(\cos qa + \frac{mV}{\hbar^2 q} \sin qa \right) e^{ika} + 1 = 0 \tag{5.39}$$

The sum of the two roots is $\cos ka$:

$$\cos ka = \frac{\cos(qa - \delta)}{\cos \delta} \tag{5.40}$$

where

$$\tan \delta = \frac{mV}{\hbar^2 q} \tag{5.41}$$

For each $k \in \left[-\frac{\pi}{a}, \frac{\pi}{a} \right]$, there are infinitely many roots q of this equation, $q_n(k)$. The energy spectrum of the nth band is:

$$E_n(k) = \frac{\hbar^2}{2m} [q_n(k)]^2 \tag{5.42}$$

$\pm k$ have the same root $q_n(k) = q_n(-k)$. Not all q's are allowed. For instance, the values $qa - \delta = n\pi$ are not allowed. These regions are the energy gaps between bands. Consider, for instance, $k = \pi/a$.

$$\cos(qa - \delta) = \cos \delta \tag{5.43}$$

This has the solutions

$$qa = \pi , \ \pi + 2\delta \tag{5.44}$$

For V small, the latter solution occurs at $qa = \pi + \frac{2mVa}{\pi\hbar^2}$. The energy gap is:

$$\begin{aligned}
E_2\left(\frac{\pi}{a}\right) - E_1\left(\frac{\pi}{a}\right) &\approx E\left(\frac{\pi}{a} + \frac{2mVa}{\pi\hbar^2}\right) - E\left(\frac{\pi}{a}\right) \\
&\approx 2V/a
\end{aligned} \tag{5.45}$$

5.5   Nearly Free Electron Approximation

According to Bloch's theorem, electronic wavefunctions can be expanded as:

$$\psi(x) = \sum_G c_{k-G}\, e^{i\left(\vec{k} - \vec{G}\right)\cdot\vec{x}} \tag{5.46}$$

In the *nearly free electron* approximation, we assume that electronic wavefunctions are given by the superposition of a small number of plane waves. This approximation is valid, for instance, when the periodic potential is weak and contains a limited number of reciprocal lattice vectors.

Let's see how this works. Schrödinger's equation in momentum space reads:

$$\left(\frac{\hbar^2 k^2}{2m} - \epsilon(\vec{k})\right) c_k + \sum_G c_{k-G}\, V_G = 0 \tag{5.47}$$

Second-order perturbation theory tells us that (let's assume that $V_k = 0$)

$$\epsilon(\vec{k}) = \epsilon_0(\vec{k}) + \sum_{G \neq 0} \frac{|V_G|^2}{\epsilon_0(\vec{k}) - \epsilon_0(\vec{k} - \vec{G})} \tag{5.48}$$

where

$$\epsilon_0(\vec{k}) = \frac{\hbar^2 k^2}{2m} \tag{5.49}$$

Perturbation theory will be valid so long as the second term is small, i.e. so long as

$$|V_G| \ll \epsilon_0(\vec{k}) - \epsilon_0(\vec{k} - \vec{G}) \tag{5.50}$$

For generic $\vec{k}$, this will be valid if V_G is small. The correction to the energy will be $O\left(|V_G|^2\right)$. However, no matter how small V_G is, perturbation theory fails for degenerate states,

$$\frac{\hbar^2 k^2}{2m} = \frac{\hbar^2 (k - G)^2}{2m} \tag{5.51}$$

or, when the Bragg condition is satisfied,

$$G^2 = 2\vec{k} \cdot \vec{G} \tag{5.52}$$

In other words, perturbation theory fails when $\vec{k}$ is near a Brillouin zone boundary.

Suppose that V_G is very small so that we can neglect it away from the Brillouin zone boundaries. Near a zone boundary, we can focus on the reciprocal lattice vector which it bisects, $\vec{G}$ and ignore $V_{G'}$ for $\vec{G} \neq \vec{G}'$. We keep only c_k and c_{k-G}, where $\epsilon_0(k) \approx \epsilon_0(k-G)$. We can thereby reduce Schrödinger's equation to a 2×2 equation:

$$
\begin{aligned}
\left(\epsilon_0(\vec{k}) - \epsilon(\vec{k})\right) c_k + c_{k-G} V_G &= 0 \\
\left(\epsilon_0(\vec{k} - \vec{G}) - \epsilon(\vec{k} - \vec{G})\right) c_{k-G} + c_k V_G^* &= 0
\end{aligned}
\tag{5.53}
$$

$V_{G'}$ for $\vec{G} \neq \vec{G}'$ can be handled by perturbation theory and, therefore, neglected in the small $V_{G'}$ limit. In this approximation, the eigenvalues are:

$$\epsilon_\pm(\vec{k}) = \frac{1}{2}\left[\epsilon_0(\vec{k}) + \epsilon_0(\vec{k} - \vec{G}) \pm \sqrt{\left(\epsilon_0(\vec{k}) - \epsilon_0(\vec{k} - \vec{G})\right)^2 + 4|V_G|^2}\right] \tag{5.54}$$

At the zone boundary, the bands have been split by

$$\epsilon_+(\vec{k}) - \epsilon_-(\vec{k}) = 2\,|V_G| \tag{5.55}$$

The effects of $V_{G'}$ for $\vec{G} \neq \vec{G}'$ are now handled perturbatively.

To summarize, the nearly free electron approximation gives energy bands which are essentially free electron bands away from the Brillouin zone boundaries; near the Brillouin zone boundaries, where the electronic crystal momenta satisfy the Bragg condition, gaps are opened.

Though intuitively appealing, the nearly free electron approximation is not very reasonable for real solids. Since

$$V_G \approx \frac{4\pi Z e^2}{G^2} \sim 13.6\,\text{eV} \tag{5.56}$$

while $\epsilon_F \sim 10 eV$,

$$|V_G| \sim \epsilon_0(\vec{k}) - \epsilon_0(\vec{k} - \vec{G}) \tag{5.57}$$

and the nearly free electron approximation is not valid.

5.6 Some General Properties of Electronic Band Structure

Much, much more can be said about electronic band structure. There are many approximate methods of obtaining energy spectra for more realistic potentials. We will content ourselves with two observations.

Band Overlap. In $2D$ and $3D$, bands can overlap in energy. As a result, both the first and second bands can be partially filled when there are two electrons per site. Consider, for instance, a weak periodic potential of rectangular symmetry:

$$V(x, y) = V_x \cos \frac{2\pi}{a_1} x + V_y \cos \frac{2\pi}{a_2} x \tag{5.58}$$

with $V_{x,y}$ very small and $a_1 \gg a_2$. Using the nearly free electron approximation, we have a spectrum which is essentially a free-electron parabola, with small gaps opening at the zone boundary. Since the Brillouin zone is much shorter in the k_x-direction, the Fermi sea will cross the zone boundary in this direction, but not in the k_y-direction. Hence, there will be empty states in the first Brillouin zone, near $(0, \pm\pi a_2)$ and occupied states in the second Brillouin zone, near $(\pm\pi a_1, 0)$. This is a general feature of $2D$ and $3D$ bands. As a result, a solid can be metallic even when it has two electrons per unit cell.

van Hove singularities. A second feature of electronic energy spectra is the existence of van Hove singularities. They are singularities in the electronic density of states, $g(\epsilon)$

$$\int d\epsilon \, g(\epsilon) \, f(\epsilon) = \int \frac{d^3k}{(2\pi)^3} \, f\left(\epsilon(k)\right) \tag{5.59}$$

They occur for precisely the same reason as in the case of phonon spectra – as a result of the lattice periodicity.

Consider the case of a tight-binding model on the square lattice with nearest-neighbor hopping only.

$$\epsilon(k) = -2t \left(\cos k_x a + \cos k_y a \right) \tag{5.60}$$

$$\vec{\nabla}_k \epsilon(k) = 2ta \left(\sin k_x a + \sin k_y a \right) \tag{5.61}$$

The density of states is given by:

$$g(\epsilon) = 2 \int \frac{d^2 k}{(2\pi)^2} \delta \left(\epsilon - \epsilon(k) \right) \tag{5.62}$$

Let's change variables in the integral on the right to E and S which is the arc length around an equal energy contour $\epsilon = \epsilon(k)$:

$$
\begin{aligned}
g(\epsilon) &= \frac{1}{2\pi^2} \int dS \frac{dE}{\left| \vec{\nabla}_k \epsilon(k) \right|} \delta \left(\epsilon - E \right) \\
&= \frac{1}{2\pi^2} \int dS \frac{1}{\left| \vec{\nabla}_k \epsilon(k) \right|}
\end{aligned}
\tag{5.63}
$$

The denominator on the right-hand-side vanishes at the minimum of the band, $\vec{k} = (0,0)$, the maxima $\vec{k} = (\pm\pi/a, \pm\pi/a)$ and the saddle points $\vec{k} = (\pm\pi/a, 0), (0, \pm\pi/a)$. At the latter points, the density of states will have divergent slope.

5.7 The Fermi Surface

The Fermi surface is defined by

$$\epsilon_n(k) = \mu \tag{5.64}$$

By the Pauli principle, it is the surface in the Brillouin zones which separates the occupied states, $\epsilon_n(k) < \mu$, inside the Fermi surface from the unoccupied states $\epsilon_n(k) > \mu$ outside the Fermi surface. All low-energy electronic excitations involve *holes* just below the Fermi surface or electrons just above it. Metals have a Fermi

surface and, therefore, low-energy excitations. Insulators have no Fermi surface: μ lies in a band gap, so there is no solution to (5.64).

In the low-density limit the Fermi surface is approximately circular (in $2D$) or spherical (in $3D$). Consider the $2D$ tight-binding model

$$\epsilon(k) = -2t \left(\cos k_x a + \cos k_y a \right) \tag{5.65}$$

For $\vec{k} \to 0$,

$$\epsilon(k) \approx -4t + ta^2 \left(k_x^2 + k_x^2 \right) \tag{5.66}$$

Hence, for $\mu + 4t \ll t$, the Fermi surface is given by the circle:

$$k_x^2 + k_x^2 = \frac{\mu + 4t}{ta^2} \tag{5.67}$$

Similarly, in the nearly free electron approximation,

$$\epsilon(\vec{k}) = \frac{\hbar^2 k^2}{2m} + \sum_{G \neq 0} \frac{|V_G|^2}{\frac{\hbar^2 k^2}{2m} - \frac{\hbar^2 (k-G)^2}{2m}} \tag{5.68}$$

For $\mu \to 0$ and V_G small, we can neglect the scond term, and, as in the free electron case, the Fermi surface is given by

$$k = \frac{1}{\hbar} \sqrt{2m\mu} \tag{5.69}$$

Away from the bottom of a band, however, the Fermi surface can look quite different. In the tight-binding model, for instance, for $\mu = 0$, the Fermi surface is the diamond $k_x \pm k_y = \pm \pi/a$.

The chemical potential at zero temperature is usually called the Fermi energy, ϵ_F. The key measure of the number of low-lying states which are available to an electronic system is the density of states at the Fermi energy, $g(\epsilon_F)$. When $g(\epsilon_F)$ is large, the C_V, σ, etc. are large; when $g(\epsilon_F)$ is small, these quantities are small.

5.8   Metals, Insulators, and Semiconductors

Earlier we saw that, in order to compute the vibrational properties of a solid, we needed to determine the phonon spectra of the crystal. A characteristic feature of these phonon spectra is that there is always an acoustic mode with $\omega(k) \sim k$ for k small. This mode is responsible for carrying sound in a solid, and it always gives a $C_V^{\mathrm{ph}} \sim T^3$ contribution to the specific heat.

In order to compute the electronic properties of a solid, we must similarly determine the electronic spectra. If we ignore the interactions between electrons, the electronic spectra are determined by the single-electron energy levels in the periodic potential due to the ions. These energy spectra break up into bands. When there is a partially filled band, there are low energy excitations, and the solid is a metal. There will be a $C_V^{\mathrm{el}} \sim T$ electronic contribution to the specific heat, as in a free fermion gas. When all bands are either filled or completely empty, there is a gap between the many-electron ground state and the first excited state; the solid is an insulator and there is a negligible contribution to the low-temperature specific heat. Let us recall how this works. Once we have determined the electronic band structure, $\epsilon_n(k)$, we can determine the electronic density-of-states:

$$g(\epsilon) = 2\sum_n \int_{\mathrm{B.Z.}} \frac{d^2k}{(2\pi)^2}\, \delta\left(\epsilon - \epsilon_n(k)\right) \tag{5.70}$$

With the density-of-states in hand, we can compute the thermodynamics. In the limit $k_B T \ll \epsilon_F$,

$$
\begin{aligned}
\frac{N}{V} &= \int_0^\infty d\epsilon\, g(\epsilon)\, \frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \int_0^\mu d\epsilon\, g(\epsilon) + \int_0^\mu d\epsilon\, g(\epsilon)\left(\frac{1}{e^{\beta(\epsilon-\mu)}+1}-1\right) + \int_\mu^\infty d\epsilon\, g(\epsilon)\, \frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&= \int_0^{\epsilon_F} d\epsilon\, g(\epsilon) + \int_{\epsilon_F}^\mu d\epsilon\, g(\epsilon) - \int_0^\mu d\epsilon\, g(\epsilon)\, \frac{1}{e^{-\beta(\epsilon-\mu)}+1} + \int_\mu^\infty d\epsilon\, g(\epsilon)\, \frac{1}{e^{\beta(\epsilon-\mu)}+1} \\
&\approx \frac{N}{V} + (\mu - \epsilon_F)\, g(\epsilon_F) + \int_0^\infty \frac{k_B T\, dx}{e^x+1}\left(g\left(\mu+k_B Tx\right) - g\left(\mu-k_B Tx\right)\right) + O\left(e^{-\beta\mu}\right)
\end{aligned}
$$

$$
\begin{aligned}
&= \frac{N}{V} + (\mu - \epsilon_F)\, g(\epsilon_F) + \sum_{n=1}^{\infty} \frac{(k_B T)^{n+1}}{n!} g^{(n)}(\mu) \int_0^{\infty} dx\, \frac{x^{2n-1}}{e^x + 1} \\
&\approx \frac{N}{V} + (\mu - \epsilon_F)\, g(\epsilon_F) + (k_B T)^2\, g'(\epsilon_F)\, I_1
\end{aligned}
\tag{5.71}
$$

with

$$
I_k = \int_0^{\infty} dx\, \frac{x^k}{e^x + 1}
\tag{5.72}
$$

We will only need

$$
I_1 = \frac{\pi^2}{6}
\tag{5.73}
$$

Hence, to lowest order in T,

$$
(\mu - \epsilon_F)\, g(\epsilon_F) \approx -(k_B T)^2\, g'(\epsilon_F)\, I_1
\tag{5.74}
$$

Meanwhile,

$$
\begin{aligned}
\frac{E}{V} &= \int_0^{\infty} d\epsilon\, \epsilon\, g(\epsilon) \frac{1}{e^{\beta(\epsilon - \mu)} + 1} \\
&= \int_0^{\epsilon_F} d\epsilon\, \epsilon\, g(\epsilon) + \int_{\epsilon_F}^{\mu} d\epsilon\, \epsilon\, g(\epsilon) + \int_0^{\mu} d\epsilon\, \epsilon\, g(\epsilon) \left(\frac{1}{e^{\beta(\epsilon - \mu)} + 1} - 1 \right) + \int_{\mu}^{\infty} d\epsilon\, \epsilon\, g(\epsilon) \frac{1}{e^{\beta(\epsilon - \mu)} + 1} \\
&\approx \frac{E_0}{V} + (\mu - \epsilon_F)\, \epsilon_F g(\epsilon_F) - \int_0^{\mu} d\epsilon\, \epsilon\, g(\epsilon) \frac{1}{e^{-\beta(\epsilon - \mu)} + 1} + \int_{\mu}^{\infty} d\epsilon\, \epsilon\, g(\epsilon) \frac{1}{e^{\beta(\epsilon - \mu)} + 1} \\
&= \frac{E_0}{V} + (\mu - \epsilon_F)\, \epsilon_F g(\epsilon_F) + \int_0^{\infty} \frac{k_B T\, dx}{e^x + 1} \Big((\mu + k_B T x)\, g(\mu + k_B T x) - \\
&\hspace{6cm} (\mu - k_B T x)\, g(\mu - k_B T x) \Big) + O\left(e^{-\beta \mu} \right) \\
&\approx \frac{E_0}{V} + (\mu - \epsilon_F)\, \epsilon_F g(\epsilon_F) + (k_B T)^2 \left[g(\epsilon_F) + \epsilon_F\, g'(\epsilon_F) \right] I_1
\end{aligned}
\tag{5.75}
$$

Substituting (5.74) into the final line of (5.75), we have:

$$
\frac{E}{V} = \frac{E_0}{V} + (k_B T)^2 g(\epsilon_F)\, I_1
\tag{5.76}
$$

Hence, the electronic contribution to the low-temperature specific heat of a crystalline solid is:

$$
\frac{C_V}{V} = \frac{\pi^2}{3}\, k_B^2 T g(\epsilon_F)
\tag{5.77}
$$

In a metal, the Fermi energy lies in some band; hence $g\left(\epsilon_F\right)$ is non-zero. In an insulator, all bands are either completely full or completely empty. Hence, the Fermi energy lies between two bands, and $g\left(\epsilon_F\right) = 0$.

Each band contains twice as many single-electron levels (the factor of 2 comes from the spin) as there are lattice sites in the solid. Hence, an insulator must have an even number of electrons per unit cell. A metal will result if there is an odd number of electrons per unit cell (unless the electron-electron interactions, which we have neglected, are strong); as a result of band overlap, a metal can also result if there is an even number of electrons per unit cell.

A semiconductor is an insulator with a small band gap. A good insulator will have a band gap of $E_g \sim 4\mathrm{eV}$. At room temperature, the number of electrons which will be excited out of the highest filled band and into the lowest empty band will be $\sim e^{-E_g/2k_BT} \sim 10^{-35}$ which is negligible. Hence, the filled and empty bands will remain filled and empty despite thermal excitation. A semiconductor can have a band gap of order $E_g \sim 0.25 - 1\mathrm{eV}$. As a result, the thermal excitation of electrons can be as high as $\sim e^{-E_g/2k_BT} \sim 10^{-2}$. Hence, there will be a small number of carriers excited into the empty band, and some conduction can occur. Doping a semiconductor with impurities can enhance this.

The basic property of a metal is that it conducts electricity. Some insight into electrical conduction can be gained from the classical equations of motion of a electron, i.e. Drude theory:

$$
\begin{aligned}
\frac{d}{dt}\vec{r} &= \frac{1}{m}\vec{p} \\
\frac{d}{dt}\vec{p} &= -eE(\vec{r},t) - \frac{e}{m}\vec{p} \times B(\vec{r},t)
\end{aligned}
\tag{5.78}
$$

If we continue to treat the electric and magnetic fields classically, but treat the electrons in a periodic potential quantum mechanically, this is replaced by:

$$
\frac{d}{dt}\vec{r} = v_n(k) = \frac{1}{\hbar}\vec{\nabla}_k \epsilon_n(k)
$$

$$\hbar \frac{d}{dt}\vec{k} \quad = \quad -eE(\vec{r},t) - e\,v_n(k) \times B(\vec{r},t) - \frac{\hbar}{\tau}\,\vec{k} \tag{5.79}$$

The final term in the second equation is the scattering rate. It is caused by effects which we have neglected in our analysis thus far: impurities, phonons, and electron-electron interactions. Without these effects, electrons would accelerate forever in a constant electric field, and the conductivity would be infinite. As a result of scattering, σ is finite. Hence, a finite electric field leads to a finite current:

$$\vec{j} = \sum_n \int \frac{d^3k}{(2\pi)^3}\,\frac{1}{\hbar}\vec{\nabla}_k\epsilon_n(k) \tag{5.80}$$

Filled bands give zero contribution to the current since they vanish by integration by parts. Since an insulator has only filled or empty bands, it cannot carry current. Hence, it is not characterized by its conductivity but, instead, by its dielectric constant, ϵ.

5.9   Electrons in a Magnetic Field: Landau Bands

In 1879, E.H. Hall performed an experiment designed to determine the sign of the current-carrying particles in metals. If we suppose that these particles have charge e (with a sign to be determined) and mass m, the classical equations of motion of charged particles in an electric field, $\mathbf{E} = E_x\hat{\mathbf{x}} + E_y\hat{\mathbf{y}}$, and a magnetic field, $\mathbf{B} = B\hat{\mathbf{z}}$ are:

$$\begin{aligned} \frac{dp_x}{dt} &= eE_x - \omega_c p_y - p_x/\tau \\ \frac{dp_y}{dt} &= eE_y + \omega_c p_x - p_y/\tau \end{aligned} \tag{5.81}$$

where $\omega_c = eB/m$ and τ is a relaxation rate determined by collisions with impurities, other electrons, etc. These are the equations which we would expect for free particles. In a crystalline solid, the momentum $\vec{p}$ must be replaced by the crystal momentum

and the velocity of an electron is no longer $\vec{p}/m$, but is, instead,

$$\vec{v}(p) = \vec{\nabla}_p \epsilon(p) \tag{5.82}$$

We won't worry about these subtleties for now. In the systems which we will be considering, the electron density will be very small. Hence, the electrons will be close to the bottom of the band, where we can approximate:

$$\epsilon(k) = \epsilon_0 + \frac{\hbar^2 k^2}{2m_b} + \dots \tag{5.83}$$

where m_b is called the *band mass*. For instance, in the square lattice nearest-neighbor tight-binding model,

$$\begin{aligned} \epsilon(k) &= -2t \left(\cos k_x a + \cos k_y a \right) \\ &\approx -4t + ta^2 k^2 + \dots \end{aligned} \tag{5.84}$$

Hence,

$$m_b = \frac{\hbar^2}{2ta^2} \tag{5.85}$$

In $GaAs$, $m_b \approx 0.07 m_e$. Once we replace the mass of the electron by the band mass, we can approximate our electrons by free electrons.

Let us, following Hall, place a wire along the $\hat{\mathbf{x}}$ direction in the above magnetic fields and run a current, j_x, through it. In the steady state, $dp_x/dt = dp_y/dt = j_y = 0$, we must have $E_x = \frac{m}{ne^2\tau} j_x$ and

$$E_y = -\frac{B}{ne} j_x = \frac{-e}{|e|} \frac{h}{e^2} \frac{\Phi/\Phi_0}{N} j_x \tag{5.86}$$

where n and N are the density and number of electrons in the wire, Φ is the magnetic flux penetrating the wire, and $\Phi_0 = h/e$ is the flux quantum. Hence, the sign of the charge carriers can be determined from a measurement of the transverse voltage in a magnetic field. Furthermore, according to (5.86), the density of charge carriers –

Figure 5.1: ρ_{xx} and ρ_{xy} vs. magnetic field, B, in the quantum Hall regime. A number of integer and fractional plateaus can be clearly seen. This data was taken at Princeton on a GaAs-AlGaAs heterostructure.

i.e. electrons – can be determined from the slope of the $\rho_{xy} = E_y/j_x$ vs B. At high temperatures, this is roughly what is observed.

In the *quantum Hall regime*, namely at low-temperatures and high magnetic fields, very different behavior is found in two-dimensional electron systems. ρ_{xy} passes through a series of plateaus, $\rho_{xy} = \frac{1}{\nu}\frac{h}{e^2}$, where ν is a rational number, at which ρ_{xx} vanishes, as may be seen in Figure 5.1. The quantization is accurate to a few parts in 10^8, making this one of the most precise measurements of the fine structure constant, $\alpha = \frac{e^2}{\hbar c}$, and, in fact, one of the highest precision experiments of any kind.

Some insight into this phenomenon can be gained by considering the quantum mechanics of a single electron in a magnetic field. Let us suppose that the electron's motion is planar and that the magnetic field is perpendicular to the plane. For now, we will assume that the electron is spin-polarized by the magnetic field and ignore the spin degree of freedom. The Hamiltonian,

$$H = \frac{1}{2m}\left(-i\hbar\nabla + e\,\mathbf{A}\right)^2 \tag{5.87}$$

takes the form of a harmonic oscillator Hamiltonian in the gauge $A_x = -By$, $A_y = 0$. (Here, and in what follows, I will take $e = |e|$; the charge of the electron is $-e$.) If we write the wavefunction $\phi(x,y) = e^{ik_x x}\,\phi(y)$, then:

$$H\psi = \left[\frac{1}{2m}\left(eB\,y + \hbar k_x\right)^2 + \frac{1}{2m}\left(-i\hbar\partial_y\right)^2\right]\phi(y)\,e^{ik_x x} \tag{5.88}$$

The energy levels $E_n = (n+\frac{1}{2})\hbar\omega_c$, called Landau levels, are highly degenerate because the energy is independent of k. To analyze this degeneracy, let us consider a system of size $L_x \times L_y$. If we assume periodic boundary conditions, then the allowed k_x

values are $2\pi n/L_x$ for integer n. The harmonic oscillator wavefunctions are centered at $y = \hbar k/(eB)$, i.e. they have spacing $y_n - y_{n-1} = h/(eBL_x)$. The number of these which will fit in L_y is $eBL_xL_y/h = BA/\Phi_0$. In other words, there are as many degenerate states in a Landau level as there are flux quanta.

It is often more convenient to work in symmetric gauge, $\mathbf{A} = \frac{1}{2}\mathbf{B} \times \mathbf{r}$ Writing $z = x + iy$, we have:

$$H = \frac{\hbar^2}{m}\left[-2\left(\partial - \frac{\bar{z}}{4\ell_0^2}\right)\left(\bar{\partial} + \frac{z}{4\ell_0^2}\right) + \frac{1}{2\ell_0^2} \right] \tag{5.89}$$

with (unnormalized) energy eigenfunctions:

$$\psi_{n,m}(z, \bar{z}) = z^m\, L_n^m(z, \bar{z})e^{-\frac{|z|^2}{4\ell_0^2}} \tag{5.90}$$

at energies $E_n = (n + \frac{1}{2})\hbar\omega_c$, where $L_n^m(z, \bar{z})$ are the Laguerre polynomials and $\ell_0 = \sqrt{\hbar/(eB)}$ is the magnetic length.

Let's concentrate on the lowest Landau level, $n = 0$. The wavefunctions in the lowest Landau level,

$$\psi_{n=0,m}(z, \bar{z}) = z^m\, e^{-\frac{|z|^2}{4\ell_0^2}} \tag{5.91}$$

are analytic functions of z multiplied by a Gaussian factor. The general lowest Landau level wavefunction can be written:

$$\psi_{n=0,m}(z, \bar{z}) = f(z)\, e^{-\frac{|z|^2}{4\ell_0^2}} \tag{5.92}$$

The state $\psi_{n=0,m}$ is concentrated on a narrow ring about the origin at radius $r_m = \ell_0\sqrt{2(m+1)}$. Suppose the electron is confined to a disc in the plane of area A. Then the highest m for which $\psi_{n=0,m}$ lies within the disc is given by $A = \pi\, r_{m_{max}}$, or, simply, $m_{max} + 1 = \Phi/\Phi_0$, where $\Phi = BA$ is the total flux. Hence, we see that in the thermodynamic limit, there are Φ/Φ_0 degenerate single-electron states in the lowest Landau level of a two-dimensional electron system penetrated by a uniform magnetic flux Φ. The higher Landau levels have the same degeneracy. Higher Landau levels

can, at a qualitative level, be thought of as copies of the lowest Landau level. The detailed structure of states in higher Landau levels is different, however.

Let us now imagine that we have not one, but many, electrons and let us ignore the interactions between these electrons. To completely fill p Landau levels, we need $N_e = p(\Phi/\Phi_0)$ electrons. Lorentz invariance tells us that if

$$n = p \frac{e^2}{h} B \tag{5.93}$$

then

$$j_x = p \frac{e^2}{h} E_y \tag{5.94}$$

i.e.

$$\sigma_{xy} = p \frac{e^2}{h} \tag{5.95}$$

The same result can be found by inverting the semi-classical resistivity matrix, and substituting this electron number.

Suppose that we fix the chemical potential, μ. As the magnetic field is varied, the energies of the Landau levels will shift relative to the chemical potential. However, so long as the chemical potential lies between two Landau levels (see figure 5.2), an integer number of Landau levels will be filled, and we expect to find the quantized Hall conductance, (5.95).

These simple considerations neglected two factors which are crucial to the observation of the quantum Hall effect, namely the effects of impurities and inter-electron interactions.[1] The integer quantum Hall effect occurs in the regime in which impurities dominate; in the fractional quantum Hall effect, interactions dominate. [2]

[1] We also ignored the effects of the ions on the electrons. The periodic potential due to the lattice has very little effect at the low densities relevant for the quantum Hall effect, except to replace the bare electron mass by the band mass. This can be quantitatively important. For instance, $m_b \simeq 0.07\, m_e$ in GaAs.

[2] The conventional measure of the purity of a quantum Hall device is the zero-field *mobility*, μ, which is defined by $\mu = \sigma/ne$, where σ is the zero-field conductivity. The integer quantum Hall effect was first observed by von Klitzing, Pepper, and Dorda in Si mosfets with mobility $\approx 10^4 \text{cm}^2/\text{Vs}$

Figure 5.2: (a) The density of states in a pure system. So long as the chemical potential lies between Landau levels, a quantized conductance is observed. (b) Hypothetical density of states in a system with impurities. The Landau levels are broadened into bands and some of the states are localized. The shaded regions denote extended states. (c) As we mention later, numerical studies indicate that the extended state(s) occur only at the center of the band.

5.9.1 The Integer Quantum Hall Effect

Let us model the effects of impurities by a random potential in which non-interacting electrons move. Clearly, such a potential will break the degeneracy of the different states in a Landau level. More worrisome, still, is the possibility that some of the states might be localized by the random potential and therefore unable to carry any current at all. As a result of impurities, the Landau levels are broadened into bands and some of the states are localized. The possible effects of impurities are summarized in the hypothetical density of states depicted in Figure 5.2.

Hence, we would be led to naively expect that the Hall conductance is less than $\frac{e^2}{h} p$ when p Landau levels are filled. In fact, this conclusion, though intuitive, is completely wrong. In a very instructive calculation (at least from a pedagogical standpoint), Prange analyzed the exactly solvable model of electrons in the lowest Landau level interacting with a single δ-function impurity. In this case, a single localized state, which carries no current, is formed. The current carried by each of the extended states is increased so as to exactly compensate for the localized state, and the conductance remains at the quantized value, $\sigma_{xy} = \frac{e^2}{h}$. This calculation gives an important hint of the robustness of the quantization, but cannot be easily generalized to the physically relevant situation in which there is a random distribution of impurities. To understand

---

while the fractional quantum Hall effect was first observed by Tsui, Störmer, and Gossard in GaAs-AlGaAs heterostructures with mobility $\approx 10^5 \mathrm{cm}^2/\mathrm{Vs}$. Today, the highest quality GaAs-AlGaAs samples have mobilities of $\approx 10^7 \mathrm{cm}^2/\mathrm{Vs}$.

Figure 5.3: (a) The Corbino annular geometry. (b) Hypothetical distribution of energy levels as a function of radial distance.

the quantization of the Hall conductance in this more general setting, we will turn to the beautiful arguments of Laughlin (and their refinement by Halperin), which relate it to gauge invariance.

Let us consider a two-dimensional electron gas confined to an annulus such that all of the impurities are confined to a smaller annulus, as shown in Figure 5.3. Since, as an experimental fact, the quantum Hall effect is independent of the shape of the sample, we can choose any geometry that we like. This one, the Corbino geometry, is particularly convenient. States at radius r will have energies similar to to those depicted in Figure 5.3.

Outside the impurity region, there will simply be a Landau level, with energies that are pushed up at the edges of the sample by the walls (or a smooth confining potential). In the impurity region, the Landau level will broaden into a band. Let us suppose that the chemical potential, μ, is above the lowest Landau level, $\mu > \hbar\omega_c/2$. Then the only states at the chemical potential are at the inner and outer edges of the annulus and, possibly, in the impurity region. Let us further assume that the states at the chemical potential in the impurity region – if there are any – are all localized.

Now, let us slowly thread a time-dependent flux $\Phi(t)$ through the center of the annulus. Locally, the associated vector potential is pure gauge. Hence, localized states, which do not wind around the annulus, are completely unaffected by the flux. Only extended states can be affected by the flux.

When an integer number of flux quanta thread the annulus, $\Phi(t) = p\Phi_0$, the flux can be gauged away everywhere in the annulus. As a result, the Hamiltonian in the annulus is gauge equivalent to the zero-flux Hamiltonian. Then, according

to the adiabatic theorem, the system will be in some eigenstate of the $\Phi(t) = 0$ Hamiltonian. In other words, the single-electron states will be unchanged. The only possible difference will be in the occupancies of the extended states near the chemical potential. Localized states are unaffected by the flux; states far from the chemical potential will be unable to make transitions to unoccupied states because the excitation energies associated with a slowly-varying flux will be too small. Hence, the only states that will be affected are the gapless states at the inner and outer edges. Since, by construction, these states are unaffected by impurities, we know how they are affected by the flux: each flux quantum removes an electron from the inner edge and adds an electron to the outer edge. Then, $\int I \, dt = e$ and $\int V \, dt = \int \frac{d\Phi}{dt} = h/e$, so:

$$I = \frac{e^2}{h} V \tag{5.96}$$

Clearly, the key assumption is that there are no extended states at the chemical potential in the impurity region. If there were – and there probably are in samples that are too dirty to exhibit the quantum Hall effect – then the above arguments break down. Numerical studies indicate that, so long as the strength of the impurity potential is small compared to $\hbar\omega_c$, extended states exist only at the center of the Landau band (see Figure 5.2). Hence, if the chemical potential is above the center of the band, the conditions of our discussion are satisfied.

The other crucial assumption, emphasized by Halperin, is that there are gapless states at the edges of the system. In the special setup which we assumed, this was guaranteed because there were no impurities at the edges. In the integer quantum Hall effect, these gapless states are a one-dimensional chiral Fermi liquid. Impurities are not expected to affect this because there can be no backscattering in a totally chiral system. More general arguments, which we will mention in the context of the fractional quantum Hall effect, relate the existence of gapless edge excitations to gauge invariance.

One might, at first, be left with the uneasy feeling that these gauge invariance arguments are somehow too 'slick.' To allay these worries, consider the annulus with a wedge cut out, which is topologically equivalent to a rectangle. In such a case, some of the Hall current will be carried by the edge states at the two cuts (*i.e.* the edges which run radially at fixed azimuthal angle). However, probes which measure the Hall voltage between the two cuts will effectively couple these two edges leading, once again, to annular topology.

Laughlin's argument for exact quantization will apply to the fractional quantum Hall effect if we can show that the clean system has a gap. Then, we can argue that for an annular setup similar to the above there are no extended states at the chemical potential except at the edge. Then, if threading q flux quanta removes p electrons from the inner edge and adds p to the outer edge, as we would expect at $\nu = p/q$, we would have $\sigma_{xy} = \frac{p}{q} \frac{e^2}{h}$.