Chapter 1.

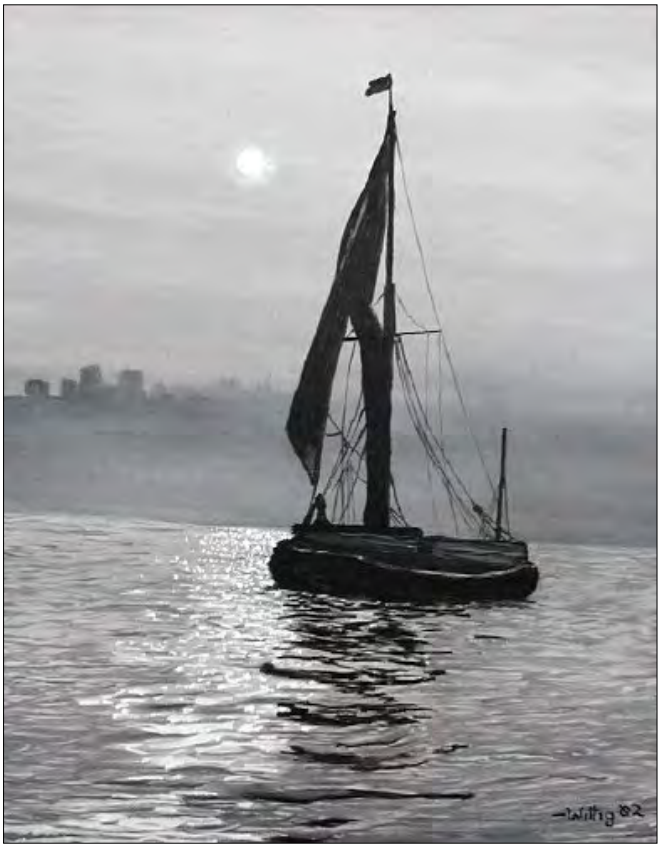
Lattice Vibrations in One Dimension

Desolation Row
Robert Wittig

Sometimes I think I'm going to end up alone in a cold apartment living on cat food from a can but I guess that's not so bad since I'm a cat after all.

Going Somewhere Soon
Brian Andreas

Contents



Barge
Robert Wittig

I pass, like night, from land to land;
I have strange powers of speech;
That moment that his face I see,
I know the man that must hear me:
To him my tale I teach.

The Ancient Mariner
Samuel Coleridge

1.1. Introduction

Nothing is so simple that it cannot be misunderstood.

This chapter introduces and examines fundamental aspects of the vibrations that take place in periodic (crystalline) structures. These vibrations, when quantized, are referred to as phonons, and the periodic structures shall be referred to hereafter as lattices. Phonons are the quanta of a lattice's vibrational normal modes. In mathematical language, they arise, in simple cases, when a field of harmonically interacting atoms that constitute a lattice is quantized using a procedure referred to as second quantization. I am not sure why this procedure is called *second* quantization – perhaps because historically the quantization of fields followed by roughly a decade the quantum mechanics of massive particles that could be treated as points, for example, electrons. Application of this second quantization approach to the lattice's mass field will be developed in a way that is hopefully easy to follow both mathematically and on physical grounds.

Like their molecular counterparts the normal modes of polyatomic molecules, phonons play a major role in a broad range of material properties. For example, were it not for phonons, electrical insulators would be perfect thermal insulators.[1] Contrast this to the fact that the thermal conductivities of many electrical insulators (MgO, BeO, Al_2O_3, diamond, and so on) are comparable to, and in some cases exceed considerably, the thermal conductivities of metals such as copper, silver, gold, aluminum, and so on.

We shall begin by considering the toy model indicated in Fig. 1 of a 1D lattice that comprises N equilibrium positions and an equal number of identical mass points. In the present chapter, 1D models will be used exclusively. They give way to 3D in Chapters 2 – 7, and to 4D relativistic spacetime in Chapters 8 and 9. In Chapters 3 – 9, we shall venture well beyond lattice vibrations: dielectric properties, bulk and surface plasmons, nanoscale plasmonics, polaritons of all kinds, and even a glimpse of relativistic field theory. We shall see that the simple model of lattice vibrations introduced here provides a foundation that facilitates understanding other field theories.

The reason for starting with the admittedly unconventional 1D lattice model indicated in Fig. 1 is pedagogical: 1D models provide straightforward, mathematically economical

[1] In the case of metals that are good thermal conductors such as copper, aluminum, silver, gold, and so on, nearly all heat transfer is brought about through scattering of phonons. Electrons play a relatively minor role because they are fermions, and therefore only a small percentage of them (those within a few kT of the Fermi surface) are free to exchange energy, and in so doing participate in heat transfer.

platforms for the introduction of concepts and phenomena. It is also a good idea to start in (or at least as close as possible to) everyone's comfort zone, and, in general, a 1D model is easier to solve mathematically than its 3D counterpart.

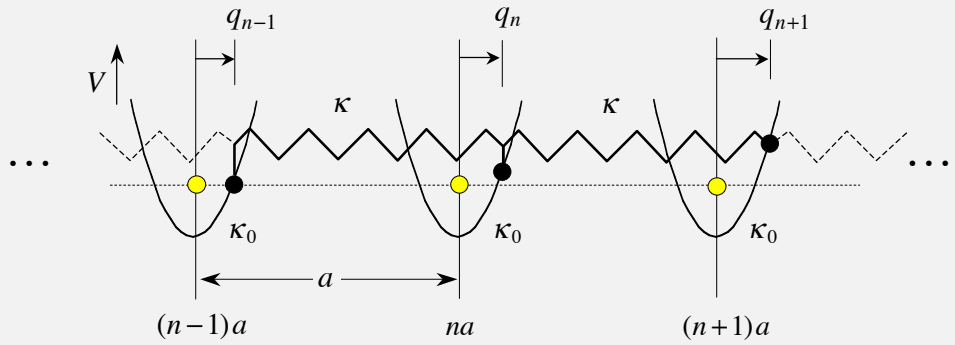


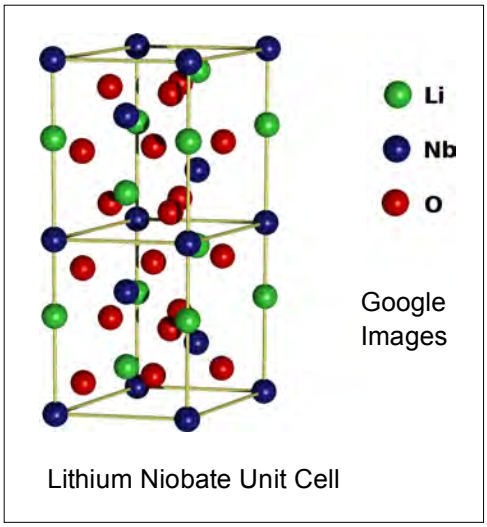
Figure 1. A one-dimensional toy model comprises identical point masses coupled to one another through identical springs and to a stationary reference frame through identical harmonic restoring forces. Black dots indicate mass points displaced from their equilibrium positions (yellow dots). The mass points are coupled to nearest neighbors with springs indicated by jagged lines. The equilibrium length of each spring is a. In addition, the mass points are coupled to their respective equilibrium positions harmonically, as indicated by the parabolic potentials. The vertical direction indicates the potential energy of the parabolas – not a spatial coordinate. The Hooke's law force constants for coupling to nearest neighbors and equilibrium positions are κ and κ_0, respectively. In many cases, κ_0 is quite a bit larger than κ.

Referring to Fig. 1, the equilibrium positions of the mass points are indicated by yellow dots spaced evenly from one another by the distance a. Black dots indicate the corresponding mass points. The mass points are bound harmonically to their respective equilibrium positions, as indicated by the parabolas. The Hooke's law force constants κ_0 for this binding to the equilibrium positions each has the same value. That is, $F = -\kappa_0 q_n$ for all n. The equilibrium positions (yellow dots) are fixed relative to a reference frame that we can think of as the laboratory.

The displacement of the n^{th} mass point from its equilibrium position is denoted q_n, and its corresponding momentum is denoted p_n. We shall ignore, for the time being, effects due to the presence of electromagnetic fields. These will be introduced later when the need arises. It is assumed that the mass points interact harmonically with one another through springs that are indicated as thick jagged lines in Fig. 1. The force constants of these springs, κ, all have the same value. Note that mass points interact directly with one another only via nearest-neighbor coupling. For example, it is assumed that the mass at q_{n-1} does not interact with the mass at q_{n+1} except via the mass at q_n. The equilibrium length of the springs is a.

The model presented in Fig. 1 is one of identical mass points and harmonic restoring forces, whereas our ultimate interest lies with lattices whose repeating units, in general, contain a number of atoms. A given repeating unit might contain several different atoms

and bonding arrangements. For example, note the primitive unit cell of a lithium niobate crystal in the sketch below. Such a crystal structure is about as complicated as I am willing to get. It is certainly quite different than the model indicated in Fig. 1.



Lithium Niobate Unit Cell

The parameters κ and κ_0 will be referred to as spring constants despite the fact that we are not really dealing with springs. For example, the Hooke's law force constant κ_0 for the parabolas in Fig. 1 shall be referred to as a spring constant to maintain a simple terminology, despite the fact that no spring is indicated in the figure. Nor could one be introduced, at least while keeping the figure simple, as the equilibrium positions are located at $q_n = 0$ for all n.

The equations of motion for the displacements can be obtained through the formal route of first writing a Lagrangian and then applying the Euler-Lagrange equation.[2] However, in the simple 1D model under consideration here, the equations of motion are written down directly by applying Hooke's law, namely, that the restoring force is proportional to the displacement. Referring to Fig. 1, applying Hooke's law to a given q_n yields

$$m\ddot{q}_n = -\kappa_0 q_n - \kappa\left(2q_n - q_{n-1} - q_{n+1}\right). \tag{1.1}$$

It would be easy to refine this equation of motion by including next-nearest-neighbor (and yet higher-order) pairwise interactions. For example, $-\kappa'(2q_n - q_{n-2} - q_{n+2})$ could be added to the right hand side of eqn (1.1), where κ' is the spring constant for springs that span two sites. To visualize this, imagine adding relatively weak springs, κ',

---

[2] The Lagrangian for the model indicated in Fig. 1 is

$$L = T - V = \sum_n \left(\tfrac{1}{2} m\dot{q}_n^2 - \tfrac{1}{2}\kappa_0 q_n^2 - \tfrac{1}{2}\kappa\left(q_n - q_{n-1}\right)^2\right).$$

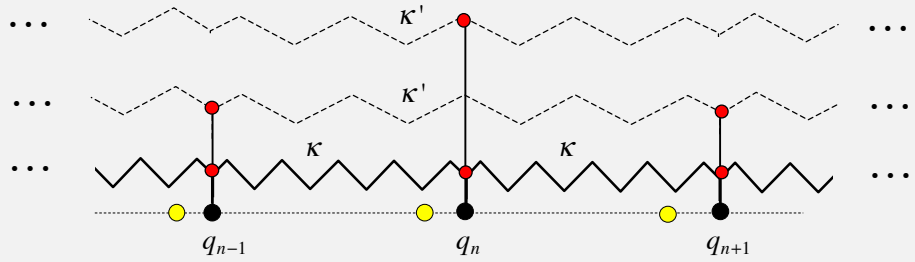
Applying the Euler-Lagrange equation to a given $q_n / \dot{q}_n$ pair yields

$$\frac{\partial L}{\partial q_n} - \frac{d}{dt}\frac{\partial L}{\partial \dot{q}_n} = 0 = -\kappa_0 q_n - \kappa\left(2q_n - q_{n-1} - q_{n+1}\right) - m\ddot{q}_n .$$

This is eqn (1.1). Note that care has been taken with the sign ambiguity that arises when taking a derivative, namely,

$$\frac{\partial}{\partial q_n}\left(q_{n+1} - q_n\right)^2 = \frac{\partial}{\partial q_n}\left(q_n - q_{n+1}\right)^2 \Rightarrow \pm 2\left(q_{n+1} - q_n\right).$$

throughout the lattice: from q_{n-1} to q_{n+1}; from q_n to q_{n+2}; and so on, as indicated in the diagram below.



The springs are displaced vertically from the mass and equilibrium points for viewing convenience. Nearest-neighbor interactions take place through springs indicated by jagged heavy black lines. Next-nearest-neighbor couplings are indicated with jagged dashed lines. Red dots indicate points where masses and springs are connected.

Anharmonicity could be included as well. However, I do not believe that additions such as next-nearest-neighbor couplings and anharmonicity yield anything of sufficient value conceptually to justify their inclusion, at least at this point in the development. Thus, we will stick to the simple expression given by eqn (1.1). Note that the Greek letter kappa is used to label spring constants. Many texts use k to label spring constants; for a harmonic oscillator one often sees $\omega^2 = k/m$. We shall use κ however, because the symbol k will be introduced later as a label for wave vectors.

The Hamiltonian for this system is given by

$$H = \sum_{n=1}^{N} \left(\frac{{p_n}^2}{2m} + \tfrac{1}{2}\kappa_0 {q_n}^2 + \tfrac{1}{2}\kappa \left(q_n - q_{n-1}\right)^2 \right). \tag{1.2}$$

It is understood that the term $\tfrac{1}{2}\kappa(q_n - q_{n-1})^2$ is equal to zero when $n = 1$. In other words, a linear open chain comprised of N sites has $N-1$ nearest-neighbor interactions.

A Nuanced Model

As mentioned earlier, the equilibrium positions represented by the yellow dots in Fig. 1 are fixed with respect to the laboratory. This differs from models in which only interparticle distances are used. When only interparticle distances appear in the equations of motion, particles can be found far from any set of laboratory-fixed reference sites, for example, a 0 K lattice, as illustrated in Fig. 2. The model introduced in Fig. 1 and eqns (1.1) and (1.2) also permits large excursions from equilibrium. However, insofar as real crystalline lattices are concerned, you will surely, and rightly so, find it counterintuitive to have the q_n displacement near the $n + 1$ site and not interact with it.

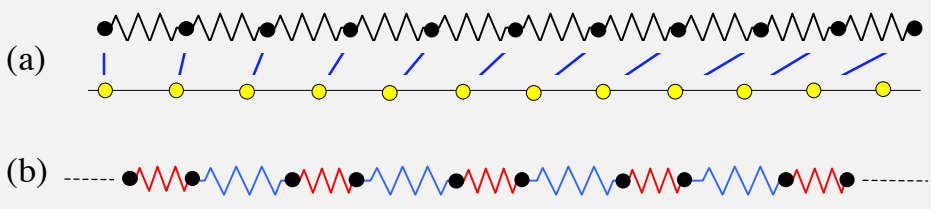


Figure 2. (a) Referring to Fig. 1, the mass points (black dots) have been displaced vertically for viewing convenience. The parabolas indicated in Fig. 1 are not shown, though they are present nonetheless. The respective equilibrium sites (yellow dots) of the mass points are indicated using short blue lines. Note that a given mass point interacts only with its own yellow dot, even when it is close to a different yellow dot. (b) Mass points are coupled to one another with springs having different spring constants, as indicated using red and blue. In (b) a laboratory reference frame plays no role.

The model indicated in Fig. 1 and eqns (1.1) and (1.2) is unusual to be sure. For example, it does not seem to represent any crystalline lattice that we have ever encountered. It couples the mass points to a fixed set of equilibrium positions in addition to coupling them to one another. In contrast, when a balls-and-springs model of a crystal is conjured, it is nearly always the case that only relative positions enter the equations of motion. The crystal as a whole, of course, sits in a laboratory reference frame. However, the atoms that constitute the crystal do not care about any laboratory reference frame.

And what about the parameter κ_0? What does it represent? If the value of the parameter κ_0 is set equal to zero, we are left with the simplest possible model: a 1D collection of identical mass points interacting via identical springs. There is no doubt that the parameter κ_0 provides an additional vibrational degree of freedom. However, it is not clear *a priori* how this additional degree of freedom relates to a 1D crystal.

Differences among lattice models are of secondary importance when introducing phonons at a qualitative, conceptual level. The important pedagogical exercises are quantization of the mass field's dynamics and variation of a given model's parameters to explore regimes and phenomena. For example, as mentioned above, the Hamiltonian for a 1D monatomic lattice composed of equal masses connected to one another by springs having equal spring constants is obtained from eqn (1.2) by setting $\kappa_0 = 0$. Though the model indicated in Fig. 1 is acceptable as a pedagogical tool, this does not explain why it was chosen. For example, later chapters deal with more realistic models of crystalline lattices, so why did we not start with one of these?

The useful feature of the toy model indicated in Fig. 1 will become clear when plasmons and polaritons are introduced and discussed in Chapters 4-7, and again in Chapters 8 and 9, where the equations of motion are assigned a covariant form. In the latter case, it turns out that the fact that the yellow dots in Fig. 1 comprise an inertial reference frame makes possible an interpretation of the model in terms of a theory in relativistic quantum field theory. Specifically, the Klein-Gordon equation is obtained through a redefinition of

the constants used in the theory. Our toy model, in fact, will prove extremely useful in Part V, Chapter 4: *Introduction to Relativistic Quantum Mechanics*. As an (optional) interlude, we shall discuss briefly the role of the parameter κ_0 from this relativistic perspective in the subsection that immediately follows this one.

In addition, many of the results obtained with the present model are applicable to the quantization of the electromagnetic field, which is discussed in Part V. This is true despite a number of important differences. The transverse photon that we are familiar with is a spin-one vector (mathematical) object, whereas phonons are spin-zero scalar objects; mass does not enter into the photon case; and the equations of motion for electromagnetic fields are Maxwell's equations, as opposed to the Schrödinger equation used here, or the Dirac equation used in the relativistic theory discussed in Part V. These differences notwithstanding, the model presented in Fig. 1 enables a great deal of material to be transported more-or-less intact to other field theories.

Coupled Wave Equations (Advanced)

This subsection provides insight regarding the planned dual-use of our toy model. The material is more advanced conceptually than one might expect for a collection of balls-and-springs. Do not be discouraged if you find it challenging. Because this subsection is self-contained, no harm is done if it is skipped or skimmed in a first reading. Keep in mind, however, that you will encounter it later, as well as additional material of the same ilk. My guess is that regardless of whether or not you skip this section now, you will come back to it in Part V. The equations in the present subsection are numbered using italic Roman numerals (*i*, *ii*, *iii*…). They do not appear outside this subsection.

To begin, eqn (1.1) is rewritten:

$$\partial_t{}^2 q_n = -\Omega_0{}^2 q_n + (\Omega a)^2 \left(\frac{(q_{n+1} - q_n) - (q_n - q_{n-1})}{a^2} \right). \qquad (i)$$

where $\partial_t \equiv \partial / \partial t$. In going from eqn (1.1) to eqn (*i*), the definitions $\kappa_0 = m\Omega_0{}^2$ and $\kappa = m\Omega^2$ have been introduced [*vide infra*, eqn (1.3)].

Notice that Ωa is the speed of sound, v. To see why this is so, set $\Omega_0 = 0$ in eqn (*i*) and assume that solutions have traveling wave character: $e^{i(kx-\omega t)}$. This assumption enables $\partial_t{}^2 q_n$ to be replaced with $-\omega^2 q_n$. The continuum limit is obtained by letting a approach zero, in which case the large parenthetic term on the right in eqn (*i*) becomes $\partial_x{}^2 q_n = -k^2 q_n$. These steps yield the elementary dispersion relation: $\omega^2 = (\Omega a)^2 k^2$. Thus, Ωa is identified as the speed of sound. Keep in mind that in taking the continuum limit it is necessary that the speed of sound and the mass density m / a remain constant.

Introducing the above facts into eqn (*i*) yields

$$(\partial_{vt}{}^2 - \partial_x{}^2)\phi = -(\Omega_0 / v)^2 \phi, \qquad (ii)$$

where $\partial_{vt} \equiv \partial / \partial(vt)$ and $\partial_x \equiv \partial / \partial x$. The mass displacement field, $\phi = (m/a)^{1/2} q_n$, replaces the discrete displacement q_n. Its form (notably the factor $(m/a)^{1/2}$) follows from the fact that integration of the Hamiltonian density must yield the same result as the one obtained using discrete mass points.[3] Factoring eqn (*ii*) and defining $\phi = \phi^{(+)}$ yields

$$\left(\partial_{vt} + \partial_x\right)\left(\partial_{vt} - \partial_x\right)\phi^{(+)} = -\left(\frac{\Omega_0}{v}\right)^2 \phi^{(+)}. \qquad (iii)$$

Next, this is split into two coupled first-order differential equations. Upon introducing the definition: $(\partial_{vt} - \partial_x)\phi^{(+)} \equiv -i(\Omega_0 / v)\phi^{(-)}$, eqn (*iii*) is converted into two first-order differential operators acting on the components, $\phi^{(+)}$ and $\phi^{(-)}$:

$$\left(\partial_{vt} + \partial_x\right)\phi^{(-)} = -i\frac{\Omega_0}{v}\phi^{(+)} \qquad (iv)$$

$$\left(\partial_{vt} - \partial_x\right)\phi^{(+)} = -i\frac{\Omega_0}{v}\phi^{(-)}, \qquad (v)$$

or, in matrix form,

$$\begin{pmatrix} 0 & \partial_{vt} + \partial_x \\ \partial_{vt} - \partial_x & 0 \end{pmatrix}\begin{pmatrix} \phi^{(+)} \\ \phi^{(-)} \end{pmatrix} = -i\frac{\Omega_0}{v}\begin{pmatrix} \phi^{(+)} \\ \phi^{(-)} \end{pmatrix}. \qquad (vi)$$

The price paid for reducing the order of the differential equation (*ii*) is that there must be two equations. The original wave function ϕ is expressed in terms of the pair of functions $\phi^{(+)}$ and $\phi^{(-)}$. We shall see that (*iv*) – (*vi*) aid in the interpretation of the parameter Ω_0.

---

[3] We shall not go into detail here, as this material will be discussed in depth later in Part IV.B and again in Part V. Briefly, when summation over n is replaced by integration over x ($na \to x$, $a \to dx$), the Hamiltonian for discrete mass points becomes a Hamiltonian density integrated over x. Not surprisingly, terms in the Hamiltonian for discrete mass points must retain their values when this Hamiltonian becomes an integration of its corresponding density over x. For example, consider the potential energy associated with the springs whose constant is κ. In the continuum limit, $a \to 0$, it becomes

$$\sum_n \tfrac{1}{2}\kappa\left(q_n - q_{n-1}\right)^2 \Rightarrow \lim_{a\to 0}\int dx\, \tfrac{1}{2}\frac{\kappa a^2}{m}\frac{m}{a}\left(\frac{\partial q_n}{\partial x}\right)^2.$$

Using $\kappa / m = \Omega^2$, the term $\kappa a^2 / m$ is identified as v^2. The ratio m/a is the 1D mass density ρ (mass per unit length). Thus, the limit yields

$$\int dx\, \tfrac{1}{2} v^2 \left(\frac{\partial\left(\rho^{1/2} q_n\right)}{\partial x}\right)^2.$$

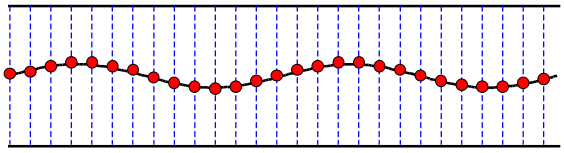
This illustrates why the displacement mass field is of the form $\phi = (m/a)^{1/2} q_n$.

When $\Omega_0 = 0$ the solutions are waves that propagate in opposite directions with phase velocity and group velocity each having magnitude v.[4] The fact that the Ω_0 term couples the waves together, and in so doing slows the speed of the resulting wave, is fascinating. We began with a discrete mass displacement field, and then took a continuum limit in which the discrete mass points were replaced with a continuous mass density. On physical grounds this sends the edge of the first Brillouin zone to infinity because the lattice no longer exists. It has been replaced with an elastic medium that supports wave motion.

Wave fronts move toward the left and right with speed v when $\Omega_0 = 0$.

In eqns (*iv*) and (*v*), one mass field acts as a source for the other and *vice versa*. In mathematical language, the Green's functions for the left hand sides of eqns (*iv*) and (*v*) are integrated over the source functions on the right hand sides. To help visualize this, consider the waves indicated in the above box, say the one moving toward the left. The Ω_0 coupling term opposes its propagation. Thus, the wave moving toward the left is undergoing continual reflection. This is the reason it appears as a source term for the wave traveling toward the right. The same idea applies to the wave moving toward the right. The take-home message is that the coupling term Ω_0 slows the velocity with which the waves travel.

Exercise: The model introduced with Fig. 1 and eqn (1) is not unique insofar as resulting in eqns $(i)-(vi)$. The underlying idea is that a wave propagating in one direction undergoes resistance in proportion to its displacement. Thus, it is a source for a wave propagating in the opposite direction. An alternate model is indicated in the sketch. Vertical springs (represented by blue dashed lines) oppose the propagation of a wave traveling to either the left or right. The wave is transverse, with vertical displacement. Show that this model also leads to eqns (*iv*) and (*v*). That is, derive equations analogous to eqns (*iv*) and (*v*).



A wave (black sinusoid) propagates to either the left or right. It is attached (red dots) to vertical springs (dashed blue lines.

With a few changes in the constants that appear in eqns (*iv*) and (*v*), the Klein-Gordon equation of relativistic quantum field theory is obtained. Specifically, the speed of sound v is replaced with the speed of light, c, and Ω_0 is replaced with $mc^2/\hbar$. Do not be alarmed by the introduction of the mass m. We are not cheating by reintroducing a mass point after eliminating discrete mass points through the transformation to a mass density. Replacing v with c automatically changes the field ϕ qualitatively: from one of mass-

[4] Taken together, eqns (*iv*) and (*v*) give the same dispersion relation as the one obtained earlier: $\omega^2 = v^2k^2$. The difference is that solutions to eqn (*ii*) with $\Omega_0 = 0$ are traveling waves with solutions that are linear combinations of waves traveling in opposite directions. In eqns (*iv*) and (*v*) the waves traveling in opposite directions are treated separately.

density displacement to one of particle-wave propagation in vacuum. The m that enters through $\Omega_0 \rightarrow mc^2/\hbar$ defines a particle.

Upon the addition of the other two spatial dimensions, eqn (*ii*) becomes

$$0 = \left(\partial_{vt}{}^2 - \nabla^2 + \left(\frac{mc}{\hbar} \right)^2 \right) \phi \,. \qquad (vii)$$

This is often written in a form that underscores its Lorentz covariant nature:

$$0 = \left(\partial^\nu \partial_\nu + \left(\frac{mc}{\hbar} \right)^2 \right) \phi \,. \qquad (viii)$$

This is the Klein-Gordon equation, albeit without the inclusion of electromagnetic fields. These are easily added once you know about gauge field theory and minimal coupling (Part V). Equations (*iv*) – (*vi*) are identical to those of *chiral waves* that are coupled to one another through the mass of the particle. All massless "particles" travel at the speed of light. Special relativity (both classical and quantum) is discussed in Part V.

Again, despite the fact that we will get to this material later, I thought it would be nice to say something here about our dual-use model. If you could follow some of the stuff in this subsection this is great. If not, do not worry.

Model Hamiltonian

Throughout the years, a number of complementary lattice models have been introduced for pedagogical purposes. Each has its merits and a few will be examined in subsequent chapters. Given the introductory nature of the material in the present chapter, qualitative understanding is a reasonable goal. Quantitative accuracy is out of the question.

Referring to eqn (1.2), the spring constants κ_0 and κ are replaced with $m\Omega_0^2$ and $m\Omega^2$, respectively, yielding

$$H = \tfrac{1}{2} \sum_{n=1}^{N} \left(\frac{p_n{}^2}{m} + m\Omega_0{}^2 q_n{}^2 + m\Omega^2 \left(q_n - q_{n-1} \right)^2 \right). \qquad (1.3)$$

In the present chapter, the parameters κ_0 and κ will no longer appear explicitly, having been replaced by the parameters $m\Omega_0{}^2$ and $m\Omega^2$, respectively. Kappa will reappear in Section 7, where the ionic lattice is discussed.

You might find it hard to believe, but an important part of the chapter is already behind us, despite the fact that we have yet to solve anything. There will be no shortage of mathematics, regimes and limiting cases will be explored, and important phenomena will be revealed. Yet, these are inevitable consequences of the ansatz that was introduced at the outset. Let us now solve the model presented in Fig. 1 and eqn (1.3).

Born and von Kármán Boundary Conditions

The system depicted in Fig. 1 has discrete translational symmetry. That is, the environment experienced by a mass point does not change from one site to the next, with the exception of the end regions. In this sense it is similar to the stationary lattices of the 1D electron models examined in Part A. In a calculation, it would be necessary to define and deal with boundary conditions at the first and last sites of a linear lattice. Intuitively, edge effects are expected to play a minor (in nearly all cases inconsequential) role for a long chain of coupled mass points. However, despite the fact that the ends are destined to play no significant role, the boundary conditions at the edges would have to be encoded into a mathematical algorithm, and this would require work that is, shall we say, of unenlightened character.

A superior approach is to impose periodic boundary conditions. They can be introduced by fiat, or through the enlistment of a physical model. For example, we can make the $n = N$ site the same as the $n = 0$ site by taking a long chain of equilibrium sites and bending it into a large circular loop, as indicated in Fig. 3. As long as N is large, there will be no significant difference in outcomes obtained using the large loop versus using the long linear chain. The loop, however, provides mathematical advantage, so it will be used. Of course, with the loop there is twofold degeneracy because waves can circulate in clockwise and counterclockwise senses.

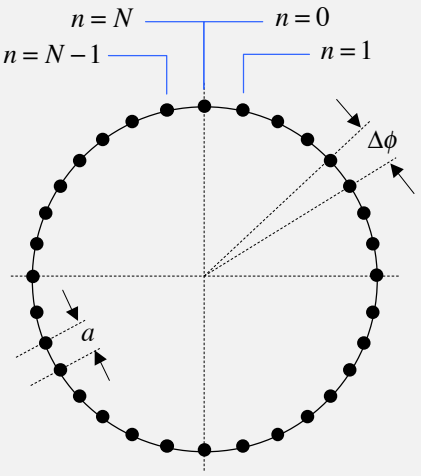


Figure 3. There are N sites; $n = 0$ and N are the same site. The angular interval $\Delta\phi$ is equal to $2\pi / N$. The distance between adjacent sites is $a = (2\pi / N) r_0$, where r_0 is the radius.

The distinction between the large loop and the linear lattice is, in many ways, the same (insofar as boundary conditions are concerned) as the distinction between a particle-on-a-ring and a particle-in-a-box. However, for the periodic lattice and large N, the difference is so small as to be immaterial. The strategy used here is identical to the one used in Part A to describe electrons in periodic 1D lattices.

With a small loop (recall benzene), the double degeneracy associated with clockwise and counterclockwise phase progression is important. However, in the present model of a large loop, this double degeneracy is not a big deal and the use of a loop causes nothing strange to happen. In the case of the loop, note that there are N nearest neighbor interactions for N sites. Thus, eqns (1.2) and (1.3) apply as written, as long as it is understood that $q_1 - q_0$ [*i.e.*, in the $n = 1$ term of the sum in eqn (1.3)] is equal to $q_1 - q_N$, as indicated in Fig. 3.

A similar situation was encountered in the Hückel model, where a matrix element β appeared in the upper right and lower left entries of the Hamiltonian matrix for a closed chain of p-orbitals. The presence or absence of β at these locations marked the mathematical distinction between closed and open chains, respectively. It was seen that the difference is dramatic for small N, but insignificant for large N.

Again, keep in mind that a lattice is not really being deformed. The loop shown in Fig. 3 is a conceptual tool enlisted for the introduction of mathematically convenient boundary conditions. Max Born and Theodore von Kármán introduced this periodic boundary condition around 1930. A curiosity is that it is not possible to extend the idea presented in Fig. 3 to a 3D lattice. The 1D lattice of Fig. 1 can be distorted into the 2D loop of Fig. 3. Likewise, a 2D lattice can be distorted first into a cylinder and then into a 3D shape called a torus, which looks like a donut or bagel. However, there is no way to distort a 3D lattice into a 3D shape that accommodates periodic boundary conditions. Topology simply does not permit it.

1.2. Site-to-Site Phase Progression

The relevant mathematical equations, both classical and quantum mechanical, can be written such that the displacements at adjacent sites differ from one another by nothing more than a phase factor $e^{\pm i\delta}$, where δ is a positive real constant. Moreover, when following a circuit around the loop, this phase factor must be the same in going from one site to the next, as the sites are identical. In other words, the phase difference between the complex displacement amplitude at the n^{th} site and the one at the $(n+1)^{th}$ site can be expressed as

$$A_{n+1} = A_n e^{-i\delta} . \tag{1.4}$$

This applies to all values of n.

There is a significant mathematical difference between classical and quantum physics when it comes to the use of complex quantities. With the former, the fundamental equations contain only real quantities. Consequently, solutions to problems in classical mechanics and classical electrodynamics, no matter how unruly the mathematics, can always be expressed as real quantities. We may choose for convenience to express them in complex form, but it is always possible to write the solutions as real quantities. For example, in classical physics complex exponentials are used frequently with the understanding that at the end of a calculation the real part is taken.[5] A big advantage of using the complex exponential form is that phases that appear in the exponent add: $e^{iA}e^{iB} = e^{i(A+B)}$, as long as A and B have a vanishing Poisson bracket.

On the other hand, quantum mechanics is built around the use of complex quantities. For example, $e^{i\vec{k}\cdot\vec{r}}$ and $e^{im\phi}$ are eigenfunctions of free particle linear momentum and the angular momentum of a particle-on-a-ring, respectively. These and other examples high-

[5] Actually, either the real or imaginary part can be taken. However, they cannot both appear in the answer.

light the fact that there is no way around the use of complex quantities; they are inescapable in quantum mechanics. Thus, when we use complex quantities in a classical calculation, it is with the understanding that the real part is (or can be) taken at the end, whereas in quantum mechanics the complex character must be retained, for example, as one sees in wave functions.

Referring to eqn (1.4), note that $|A_{n+1}| = |A_n|$. That is, the magnitudes of the complex displacement amplitudes are the same at all sites. Equation (1.4) follows from the fact that the sites are equivalent, so there is no reason to single out any one site, or group of sites, as privileged. Because of the twofold degeneracy (clockwise and counterclockwise phase progression), in the quantum mechanical case degenerate pairs of functions can be combined to create probability density that differs from site to site. The phase relationship in eqn (1.4) is essentially the same as the ones used in Part A. For example, in the Hückel model an analogous phase progression was assigned to *p*-orbitals. There, we saw that linear combinations of the degenerate partners yielded the real functions most often found in texts, a classic example being benzene. In the present case, an arbitrary complex displacement A_n with constant site-to-site magnitude is considered. It will be seen that eqn (1.4) leads to compression/rarefaction (acoustic) waves along the 1D lattice.

The choice of $-i\delta$ instead of $+i\delta$ in eqn (1.4) is arbitrary. The ± signs in $e^{\pm i\delta}$ indicate the sense of phase progression, which cannot be important. In other words, there is nothing special about the sense of phase progression when following a circuit around the loop. It is only necessary that the site-to-site phase shift retain its value everywhere on the loop. The use of $e^{-i\delta}$ will prove convenient later when other conventions are encountered, for example, when the wave vector k is introduced. This is the reason $-i\delta$ is used in eqn (1.4).

The total phase change upon the completion of a circuit around the loop is $\pm N\delta$, depending on whether the circuit follows a clockwise or counterclockwise path. This must be equal to an integer multiple of 2π. Thus, the parameter δ can be expressed as

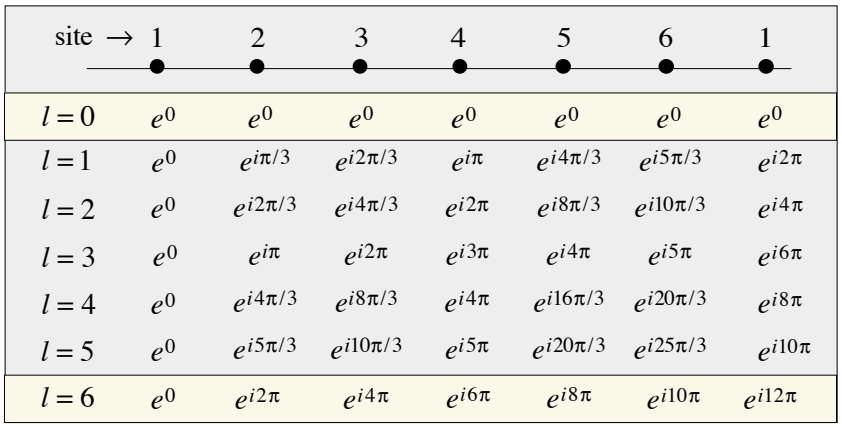
$$\delta = \frac{2\pi}{N} l \,, \tag{1.5}$$

where l is an integer that has N independent values. The reason there are only N independent values of l is uniqueness. For example, if $l = N + 1$, then $\delta = (N+1)(2\pi/N) = 2\pi + 2\pi/N$, which is equivalent (modulo 2π) to $l = 1$. Note that there is latitude in the choice of l values, as long as the number of independent l values is equal to N. For example, the range $1 \rightarrow N$ works, but so do other ranges and collections of l values.[6] We will see that the most convenient, and by far the most widely used, convention is to have l assume positive and negative values in a range centered at $l = 0$.

---

[6] On mathematical grounds, the values of l need not be successive integers. For example, if $N = 6$, the values 1, 2, 4, 6, 9, and 11 work, because $9 - 6 = 3$ and $11 - 6 = 5$. The l values must be such that they can be converted to a sequence of successive integers. This mathematical fact (for our purposes, a curiosity) does not arise in anything we will encounter. The reason it is even mentioned here is that in the past a few students have asked about it.

In examining the phase factor $e^{-i\delta}$, it is clear that $2\pi s$ (where s is an integer) can be added to δ without altering the phase in a meaningful way. This is because $e^{\pm i2\pi s} = 1$ for integer s. For example, $e^{-i\pi/4}$ and $e^{i7\pi/4}$ are, for all practical purposes, the same.[7] Figure 4 illustrates this point further. Figure 5 does likewise, showing that the same phase is obtained with the more rapidly oscillating sinusoid. In other words, because the lattice is discrete no additional information is introduced when an integer number of wavelengths is added to each of the intervals between the mass points.



$l = 0$	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}	e^{0}
$l = 1$	e^{0}	$e^{i\pi/3}$	$e^{i2\pi/3}$	$e^{i\pi}$	$e^{i4\pi/3}$	$e^{i5\pi/3}$	$e^{i2\pi}$
$l = 2$	e^{0}	$e^{i2\pi/3}$	$e^{i4\pi/3}$	$e^{i2\pi}$	$e^{i8\pi/3}$	$e^{i10\pi/3}$	$e^{i4\pi}$
$l = 3$	e^{0}	$e^{i\pi}$	$e^{i2\pi}$	$e^{i3\pi}$	$e^{i4\pi}$	$e^{i5\pi}$	$e^{i6\pi}$
$l = 4$	e^{0}	$e^{i4\pi/3}$	$e^{i8\pi/3}$	$e^{i4\pi}$	$e^{i16\pi/3}$	$e^{i20\pi/3}$	$e^{i8\pi}$
$l = 5$	e^{0}	$e^{i5\pi/3}$	$e^{i10\pi/3}$	$e^{i5\pi}$	$e^{i20\pi/3}$	$e^{i25\pi/3}$	$e^{i10\pi}$
$l = 6$	e^{0}	$e^{i2\pi}$	$e^{i4\pi}$	$e^{i6\pi}$	$e^{i8\pi}$	$e^{i10\pi}$	$e^{i12\pi}$

Figure 4. This is taken from Part A (Hückel model): Consider a circular loop (Fig. 3) and $N = 6$. Phase factors showing phase progression $e^{i(2\pi/N)l}$ are given for different l values. The phase at site 1 is set to zero. For $l = 0-5$, each of the phase progressions is distinct. However, for $l = 6$, the phases at the different sites are the same (modulo 2π) as with $l = 0$, so $l = 0$ and 6 represent the same state. Likewise, $l = 7$ and $l = 1$ are equivalent, and so on.

[7] Qualifiers such as "in a meaningful way" and "for all practical purposes" are used because in general it is necessary to include spin. For odd-half-integer spin (odd-half-integer angular momentum in general) the appropriate interval is 4π, as spinors are single-valued in 4π rather than 2π. In the present context, spin is not dealt with, so 2π is the appropriate interval.

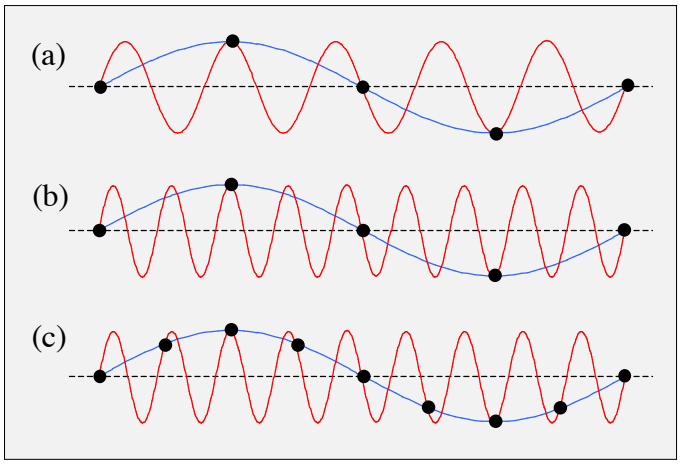


Figure 5. Transverse waves (propagation in the horizontal direction) are used to illustrate the fact that adding one or more reciprocal lattice vectors (*i.e.*, $n2\pi/a$) to k has no effect on the displacements (equivalently, on the phase, modulo 2π). (a) The blue curve gives the displacements of the black dots. The red curve does the same. No information has been added; an additional wavelength has simply been squeezed into the intervals between the dots. In other words $2\pi/a$ has been added to k. (b) The addition of $4\pi/a$ to k is shown, *i.e.*, two wavelengths have been squeezed in between dots, again leaving the displacement unaffected. (c) One wavelength is added to a denser lattice.

Figures 4 and 5 have counterparts in Part A, and similar figures are found in books on solid-state physics. In the language used to describe electrons in lattices, each Brillouin zone contains a fixed number of unique solutions, namely, the number of sites. In the case of phonons, everything happens in the first Brillouin zone. That is, δ values contained within a single 2π interval provide all of the unique phases. With electrons, though different Brillouin zones are associated with corresponding electron orbitals, it is possible to "move the Brillouin zones over" by a reciprocal lattice vector, $G = n(2\pi/a)$, in order to represent everything in the first Brillouin zone. In the present chapter, we will see that the quantized lattice vibrational modes (phonons) are represented using *just* the first Brillouin zone.

When a reciprocal lattice vector is added to the lattice, the entire lattice acquires a linear momentum. Such momentum arises in a number of inelastic scattering processes. For example, it arises in the context of thermal conductivity, as discussed in Chapter 2. However, the momentum of the reciprocal lattice vector results in an extremely small amount of energy being imparted to the lattice. To see just how small an amount of energy this might be, take $\delta E = (\hbar G)^2 / 2M$, where M is the mass of the lattice (crystal), and insert some numbers. For n = 1, a = 0.3 nm, and M = 10 μg (a quite small crystal), δE is of order 10^{-40} J. Clearly this can be neglected.

Following an almost universal convention in the literature, the 2π range of δ values is taken as:[8]

$$-\pi < \delta \le \pi \,. \tag{1.6}$$

[8] The inequality < is used on one side, whereas ≤ is used on the other. Unless the number of sites is modest, there is no significant difference between using < versus using ≤. However, for a small number of sites, the difference is profound, as we have seen in Part A.

The interval given by eqn (1.6) is referred to as the first Brillouin zone. As mentioned earlier, when there are many sites, the distinction between the range $-\pi < \delta \leq \pi$ given by eqn (1.6) and the range $-\pi \leq \delta \leq \pi$ is unimportant. In other words, we need not be careful with < versus ≤.

Now consider an ensemble of complex displacement amplitudes A_n given by eqn (1.4). To take advantage of the discrete (site-to-site) translational symmetry, this ensemble is written as a linear combination of the A_n values associated with the different sites. What we are going to do at this point is enlist Fourier analysis, which is a natural choice given the lattice periodicity. It is important that each of the complex amplitudes differs from its adjacent partner only by a constant phase factor that satisfies the condition that the total phase for a closed circuit around the loop is divided into *N* equal pieces. This is the simplest function that transforms according to the rule that if all the sites are renumbered ($1 \rightarrow 2$, $2 \rightarrow 3$, and so on), the quantum mechanical wave functions are recovered to within an overall constant phase factors, because there can be no energy difference. Let us call this linear combination Q_δ. Moreover, let us choose this linear combination such that it results in a single value of the phase parameter δ. This requires that all values of *n* are included in the expansion:

$$Q_\delta = N^{-1/2} \sum_{n=1}^{N} A_n \ .$$

$$= N^{-1/2} \sum_{n=1}^{N} e^{-i\delta n} q_n \ . \tag{1.7}$$

In going from the first to the second of these equations, the phase of A_n is appended to the real displacement q_n, *i.e.*, $A_n = e^{-i\delta n} q_n = e^{-ikan} q_n$ [*vide infra,* eqn (1.8)], which clearly satisfies eqn (1.4): $A_{n+1} = A_n e^{-i\delta}$. This enables us to work with complex amplitudes A_n and take the real part at the end of a calculation. The factor $N^{-1/2}$ takes care of normalization, which does not mean much at the moment, but whose significance materializes later.

The choice of real q_n is made in anticipation of q_n becoming a Hermitian operator when the system is made quantum mechanical. Now is as good a time as any to make the transition from phase parameter δ to wave vector *k*. Recalling eqn (1.5): $\delta = (2\pi / N) l$, this is achieved through the same rearrangement that was used in Chapter 2 of Part A.

$$\delta = \frac{2\pi}{N} l = \frac{2\pi}{Na} la = \left(\frac{2\pi}{L} l \right) a = ka \ , \tag{1.8}$$

where *a* is the distance between adjacent sites, and *l* assumes integer values. The fact that the overall length of a circuit around the loop shown in Fig. 3 is given by $L = Na$ has been used. Thus, *k* is given by

$$k = \frac{2\pi}{L} l \,. \tag{1.9}$$

The allowed k values are given by the product of the fundamental wave vector interval, $2\pi / L$, times the integer l. The integer l can assume N independent values.

When $l = \pm 1$, the wavelength, $2\pi / |k|$, is equal to the circumference L of the loop in Fig. 3. At the other extreme, when $l = N/2$ (*i.e.*, $k = \pi / a$), a half wavelength matches the lattice spacing. It is good to keep these limiting cases in mind. When $l = 0$, one could say that the wavelength is infinity, as there is no site-to-site phase shift. This case corresponds to overall displacement of the linear lattice that was deformed into a loop for mathematical convenience. No energy is required to displace the linear lattice, so $l = 0$ is unimportant, to say nothing of the fact that the number of allowed l values is very large, in which case the single $l = 0$ contribution is insignificant on statistical grounds.

Equation (1.7) is now rewritten using k instead of δ. This transforms Q_δ into Q_k. Also, the square root of the particle mass m is introduced into the definition of Q_k. The reason for doing this is most likely not obvious at this point. It turns out that this inclusion simplifies much of the math that follows. The resulting expression for Q_k is

$$Q_k = \sqrt{\frac{m}{N}} \sum_n e^{-ikan} q_n \,, \tag{1.10}$$

where $k = (2\pi / L) l$. Two integers are present in eqn (1.10): n and l. The former indexes the lattice sites, while the latter quantizes k.

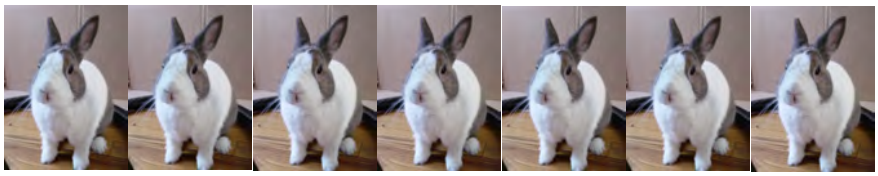
Equation (1.10) is a Fourier expansion, with Q_k expressed in terms of phase factors e^{-ikan} and displacements q_n for the different sites. For a specific value of k, the displacements q_n that describe Q_k span the entire range of n values. In other words, for a specific value of k the entire lattice is involved. So far everything is classical. Nonetheless, we see, even at this early stage, that it will not be possible for Q_k to become a Hermitian operator when the model is made quantum mechanical. The reason is that Q_k includes all of the sites, with phase attached to each site. As such, it is not an observable, so it will not become a Hermitian operator. Whereas q_n can be made real, Q_k cannot, because of the phase factors. Thus, on mathematical grounds alone, q_n and Q_k cannot each be real.

Because of the symmetry of the loop, it follows that solutions must exist for which there is no change other than phase upon progressing from site to site. Recall that this was also encountered in Part A, where we found electron orbitals that differed by a phase factor, with combinations of $+k$ and $-k$ yielding real wave functions. Again, keep in mind the distinction between phase factors that appear in classical versus quantum mechanical models. The former are a convenience, whereas the latter have physical consequence.

The same idea applies here, though the q_n in eqn (1.10) are not solutions to some zero-order dynamics problem, as were the p-orbitals in the Hückel model. Together, q_n and its counterpart momentum p_n constitute a pair of canonically conjugate variables for the n^{th} lattice site. When the lattice problem is solved quantum mechanically in coordinate space, spatial wave functions are found for each of the mass points. These wave

functions differ from one site to the next by phase factors e^{ika}. This aspect is similar to the electron wave functions in the periodic potential $V(\phi)$ in Part A, with the understanding that the many-body electron wave function must be antisymmetric with respect to exchange of any two electrons.

This relationship, in which the narrowest possible distribution of k values requires the largest possible distribution of n values, is reminiscent of the $\Delta x\,\Delta p$ uncertainty product of quantum mechanics. However, the uncertainty relation $\Delta x\,\Delta p \geq \hbar/2$ is derived on the basis of the commutator relation $[x,p]=i\hbar$, whereas here the relationship between the distributions of Q_k and q_n values is based on Fourier analysis. Keep in mind that in the present case the momentum p is not equal to $\hbar k$, as it would be for a free particle. Momentum is not a conserved quantity in a periodic potential.[9] This is because the particle's motion is affected through its interaction with the site potentials. Stated more elegantly: only a continuous symmetry can yield a conserved quantity. In other words, momentum is not conserved when the symmetry is discrete rather than continuous. This is the essence, and one of the simplest examples, of an important theorem due to the brilliant mathematician Emmy Noether. Not surprisingly, it is referred to as Noether's theorem.



Indiscreet Symmetry

Equation (1.10) is easily inverted to yield q_n as an expansion over the Q_k basis. The entire collection of displacements (q_n for all n values) is related to the entire collection of Q_k (all k values) through Fourier analysis. What we have is a discrete Fourier transform, in which the transform pair is q_n and Q_k, as opposed to the Fourier transform pair x and k, for example, that you have encountered in introductory quantum mechanics. Because Fourier analysis is remarkably well established, we have at our beckon a ready prescription.[10] Inversion is achieved by starting with $(mN)^{-1/2}$ times a sum over Q_k with accompanying phase factors. It is easily achieved as follows:

$$\frac{1}{\sqrt{mN}}\sum_k Q_k e^{ikan} = \frac{1}{\sqrt{mN}}\sum_k \left(\sqrt{\frac{m}{N}}\sum_{n'} e^{-ikan'}q_{n'}\right)e^{ikan} \tag{1.11}$$

---

[9] Strictly speaking, in the case of the loop we should talk about *angular* momentum. For a large loop, there is no significant difference, in the sense that the momentum in the vicinity of a site can be taken as linear. Moreover, we are not really interested in the loop. The periodic boundary condition could just as easily have been introduced by fiat while keeping the 1D lattice linear.

[10] Elementary Fourier analysis is covered nicely in Kusse and Westwig: *Mathematical Physics: Applied Mathematics for Scientists and Engineers* [36].

$$= \frac{1}{N}\sum_{n'} q_{n'}\left(\sum_{k} e^{ika(n-n')}\right). \tag{1.12}$$

The parenthetic term is equal to N times the Kronecker delta, $\delta_{n,n'}$, yielding

$$\frac{1}{\sqrt{mN}}\sum_{k} Q_k e^{ikan} = \frac{1}{N}\sum_{n'} q_{n'} N\delta_{n,n'}\,. \tag{1.13}$$

The N's on the right hand side cancel, and the Kronecker delta $\delta_{n,n'}$ selects the n^{th} term in the sum. Therefore the right hand side of eqn (1.13) is equal to q_n and we have

$$\boxed{q_n = \frac{1}{\sqrt{mN}}\sum_{k} Q_k e^{ikan}\,.} \tag{1.14}$$

In going from eqn (1.12) to (1.13), use has been made of the fact that the sum over k vanishes unless $n' = n$, in which case the sum over k is equal to N. This sum and similar ones come up again and again, so you might as well get used to them. They are orthogonality relations. What I am going to do in the next few pages is present three alternate ways to see how this works. They approach the issue from complementary perspectives, but they are equivalent in content. The goal is an intuitive grasp of this orthogonality condition.

Diagram from the Hückel Model

In addition to proving the orthogonality relationship, namely, that the sum over k in eqn (1.12) yields $N\delta_{n,n'}$, through the use of algebra, it is even more instructive to use a diagram. Figure 6 is taken from Part A, Chapter 3.

In the present context, the angle θ is equal to ka. For example, for N = 6, $ka = (2\pi/N)l = (\pi/3)l$, with l going from –2 to +3.[11] It is easy to see, by visual inspection or simple trigonometry, that a sum over k gives projections onto vertical and horizontal axes (whose origins are at the centers of the circles) that add to zero. With eqn (1.12) we are dealing with $ka(n-n') = (2\pi/N)l(n-n')$. When this is the argument of the complex exponential that is summed over the first Brillouin zone, a result of zero is always obtained (for $l \neq 0$) unless $n-n' = 0$ is satisfied. Thus, Fig. 6 applies.

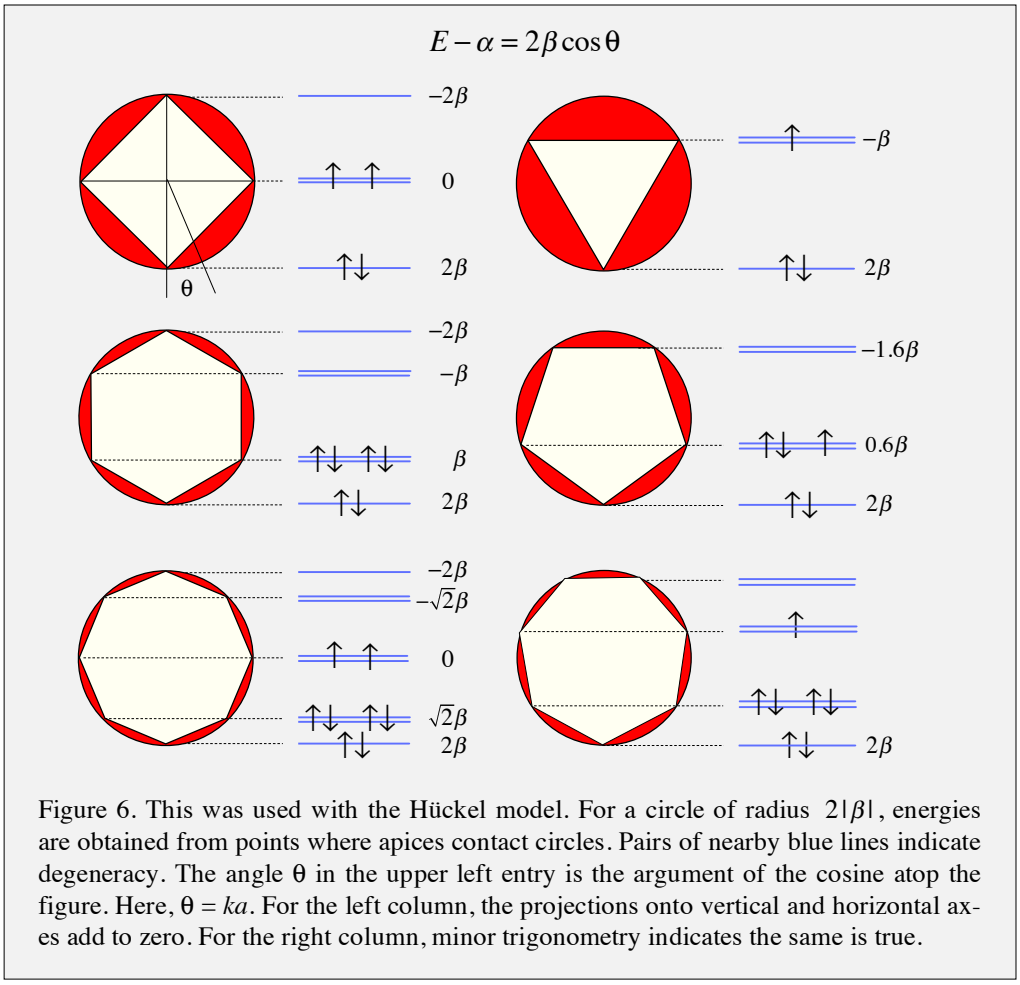


Figure 6. This was used with the Hückel model. For a circle of radius $2|\beta|$, energies are obtained from points where apices contact circles. Pairs of nearby blue lines indicate degeneracy. The angle θ in the upper left entry is the argument of the cosine atop the figure. Here, θ = ka. For the left column, the projections onto vertical and horizontal axes add to zero. For the right column, minor trigonometry indicates the same is true.

[11] Equivalently, l can go from –3 to +2, or the two l values –3 to +3 can be used with the understanding that –3 and +3 together contribute just one state.

Mathematical Approach from Phased Antenna Arrays

The Kronecker delta orthogonality relationship can also be obtained mathematically in a straightforward manner. Referring to eqn (1.12), we replace ka with $(2\pi/N)l$, in which case the parenthetic term is written

$$\sum_l e^{i(2\pi n/N)l} \ .$$

Keep in mind that there are N independent values of both n and l.

Here, we have used n instead of $n-n'$. Later it will be pointed out that the latter gives the same result. The manipulations are easy, but you might wonder how one figures out beforehand how to proceed. It turns out that this orthogonality relation is obtained via a standard mathematical trick used in other areas of science and engineering. I first encountered it with antenna radiation patterns, but it has widespread application. Anyway, let us go through the steps.

The sum defined below and labeled with capital theta is introduced for convenience. It is simplest to sum from $l = 0$ to $N - 1$.

$$\Theta = \sum_{l=0}^{N-1} e^{i(2\pi n/N)l} \tag{1.15}$$

$$= 1 + e^{i(2\pi n/N)} + e^{i2(2\pi n/N)} \ldots e^{i(N-1)(2\pi n/N)}$$

$$= 1 + e^{i(2\pi n/N)}(1 + e^{i(2\pi n/N)} \ldots e^{i(N-2)(2\pi n/N)}) + e^{iN(2\pi n/N)} - e^{iN(2\pi n/N)} \tag{1.16}$$

$$= \Theta e^{i(2\pi n/N} + (1 - e^{i2\pi n}) \ . \tag{1.17}$$

Collecting the Θ terms yields

$$\Theta(1 - e^{i2\pi n/N}) = (1 - e^{i2\pi n}) \ . \tag{1.18}$$

Thus, the desired sum is expressed as a ratio:

$$\sum_{l=0}^{N-1} e^{i(2\pi n/N)l} = \frac{1 - e^{i2\pi n}}{1 - e^{i2\pi n/N)}} \ . \tag{1.19}$$

The numerator vanishes for all integer n, whereas the denominator does not vanish for $1 \le n \le N - 1$. For $n = 0$, the denominator vanishes, so in this case the ratio is obtained using l'Hôpital's rule. This yields N for the right hand side. Notice that when n is replaced by $n - n'$ the math is the same. Even though the mathematical approach given by eqns (1.15) – (1.19) is clever, Figs. 6 and 7 are probably easier to remember.

The Complex Plane

Another – and perhaps the easiest – way to grasp the result is through the diagram in Fig. 7. The vectors in rows (a) and (b) (which represent even and odd numbers of sites, respectively) add to zero for each row. Any time 2π is divided into equal segments, the sum of complex unit vectors in a complete circuit yields zero. You should verify this by trying a few.

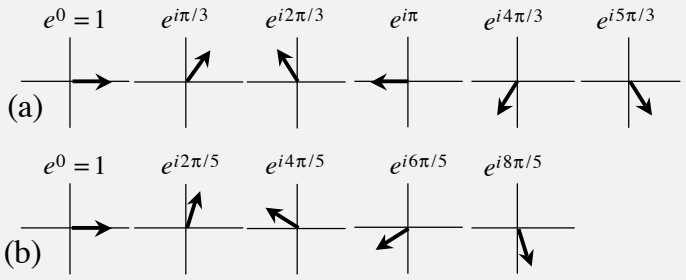


Figure 7. Phase factors represented in the complex plane: The vector sums are zero for both even and odd numbers of sites, as indicated in (a) and (b), respectively.

Physical Space and k-Space

To further interpret eqn (1.14), note that all possible wave vectors are included in the Fourier expansion of the displacement q_n. This is due to the fact that q_n is local to a single site – the n^{th} one. Fourier analysis yields the intuitive result that spatial localization is achieved by using the largest range of k values. As a matter of convenience, one k value can be set equal to zero, and the others can be arranged in pairs in which each k value has the same magnitude: $k = \pm |k|$. What is less obvious (or perhaps not obvious at all) is that Fourier analysis requires that waves traveling in opposite directions must each be present in a packet if a high degree of spatial localization is to be achieved. Indeed, it is a mathematical fact that waves traveling in opposite directions must each be present in a packet if the packet is to vanish outside a given desired region of localization.

The manipulations that were carried out in the previous subsections involved nothing more complicated than Fourier analysis. They showed how a spatial displacement q_n is expanded in the Q_k basis. The next goal is to introduce this expansion into the Hamiltonian. As mentioned earlier, transforming the Hamiltonian to its k-space representation will result in a great deal of simplification.[12] Likewise, p_n is expanded in the P_k basis and this is also introduced into the Hamiltonian. The math for the p_n expansion is analogous to the math for the q_n expansion, but with a notable difference.

The Fourier expansions for p_n and P_k are given by

[12] Intuitively, k labels phase progression, so degenerate pairs of frequencies are expected, with a sum of harmonic oscillators for the Hamiltonian.

$$p_n = \sqrt{\frac{m}{N}} \sum_k P_k e^{-ikan} \tag{1.20}$$

and

$$P_k = \frac{1}{\sqrt{mN}} \sum_n p_n e^{ikan} . \tag{1.21}$$

From before

$$q_n = \frac{1}{\sqrt{mN}} \sum_k Q_k e^{ikan} \tag{1.14}$$

$$Q_k = \sqrt{\frac{m}{N}} \sum_n q_n e^{-ikan} \tag{1.10}$$

One sees right away that the signs in the exponents are not the same with the p_n / P_k pair as with its q_n / Q_k counterpart. The box on the right has been added to facilitate this comparison.

The signs that appear in the exponents of eqns (1.20) and (1.21) are reversed from those of their q_n and Q_k counterparts, as seen in the box entries. Specifically, the summations in eqns (1.20) and (1.21) contain, respectively, $P_k e^{-ikan}$ rather than $P_k e^{ikan}$, and $p_n e^{ikan}$ rather than $p_n e^{-ikan}$. In other words, the expansion of q_n given by eqn (1.14) sums the terms $Q_k e^{ikan}$ over all k, whereas the expansion of p_n given by eqn (1.20) sums the terms $P_k e^{-ikan}$ over all k. Likewise for the inverse transformations. The expansion of Q_k given by eqn (1.10) sums the terms $q_n e^{-ikan}$ over all n, whereas the expansion of P_k given by eqn (1.21) sums the terms $p_n e^{ikan}$ over all n. This invites reasonable questions: What dictates these phase choices, in which a positive sign in one argument (for example, $Q_k e^{ikan}$) is accompanied by a negative sign in the complementary argument ($P_k e^{-ikan}$), and are these phase choices significant on physical grounds or mere conventions?

Choices of Signs

Motivated by questions, comments, and suggestions made by Vyaas and others in class (2012), and the interesting and useful discussions that ensued, it is appropriate that we look further into the choice of signs in the exponents of eqns (1.20) and (1.21). On strictly mathematical grounds, the Fourier pairs q_n / Q_k and p_n / P_k, when judged separately, can go either way insofar as signs in exponents is concerned. For example, exchanging plus and minus in eqns (1.20) and (1.21) yields an equally valid Fourier pair. After all, mathematically, the various symbols do not carry any significance. They can be interchanged with no harm done to the math. However, we shall see that when taking the physics into account, an unambiguous relationship between the phase progressions of the q_n / Q_k and p_n / P_k Fourier transform pairs is established. It is found that the choice given by eqns (1.20) and (1.21) is not simply one of two equally acceptable options. It is required.

The physics that dictates this choice is embodied in the Poisson bracket of classical mechanics and the commutator of quantum mechanics. To see how this works, we shall consider first the quantum mechanical case. Commutator expressions for n-space and k-

space will be written in terms of the respective Fourier expansions of their contents, and the interplay between the aforementioned choice of phase and consistent physics will be determined. Throughout this exercise, carats will be placed atop operators to underscore their quantum character. Upon completion of the exercise the carats are dropped, and the analogous steps for the classical case (Poisson bracket) are carried out. This will demonstrate that the choice of phase progression in the Fourier expansions is consistent with, but not *due to*, quantum mechanics. In other words, it is present in classical physics. This is not surprising, as $\hbar$ is nowhere to be found in any of the expressions derived so far.

Quantum Route

Starting with the commutator $[\hat{q}_n, \hat{p}_n]$, the expansion of $\hat{q}_n$ given by eqn (1.14) is inserted, and an analogous expansion of $\hat{p}_n$ is needed. The one given by eqn (1.20) could be used, and in this case everything would fall neatly into place. However, instead of doing this, let us carry out the math using the reversed sign in the exponent, just to see what happens. Specifically, we shall use $\hat{P}_k e^{ikan}$ instead of the $\hat{P}_k e^{-ikan}$ used in eqn (1.20). This will reveal whether each choice is acceptable, and if not, what is wrong with the alternate one under consideration below.

Pursuant to the above strategy, the commutator $[\hat{q}_n, \hat{p}_n]$ is expressed as

$$\left[\hat{q}_n, \hat{p}_n\right] = \frac{1}{N}\sum_{k,k'}\left[\hat{Q}_k e^{ikan}, \hat{P}_{k'} e^{ik'an}\right] \tag{1.22}$$

$$= \frac{1}{N}\sum_{k,k'}\left[\hat{Q}_k, \hat{P}_{k'}\right] e^{i(k+k')an} . \tag{1.23}$$

The value of the commutator $[\hat{Q}_k, \hat{P}_{k'}]$ has not yet been established. However, we know that $k = k'$ must hold if $[\hat{Q}_k, \hat{P}_{k'}]$ is to have a nonzero value, because different k values label the different k-space "sites." In the case of the mass points and their momenta (where different n values label the n-space sites), a commutator relation cannot have a nonzero value for displacement at one site and momentum at another. The same restriction applies in k-space: each k-space site has its own commutator relation. The resulting $k = k'$ requirement reduces the sum over k and k' to one over k. Thus, eqn (1.23) becomes

$$\left[\hat{q}_n, \hat{p}_n\right] = \frac{1}{N}\sum_{k}\left[\hat{Q}_k, \hat{P}_k\right] e^{i2kan} . \tag{1.24}$$

The left hand side is equal to $i\hbar$. The commutator $[\hat{Q}_k, \hat{P}_k]$ must have the same value for each k. This value can be moved to the left of the summation, leaving a sum of exponentials that vanishes identically. The right hand side being zero, with the left hand side being equal to $i\hbar$, is a dramatic inconsistency. Thus, the chosen phase progression is illegal. Had the sign given in eqn (1.20) been used instead, a perfectly acceptable result would have been obtained. In this case, tracing the steps from eqn (1.22) to (1.24) would have yielded $e^0 = 1$ instead of e^{i2kan}. In this case we have $[\hat{Q}_k, \hat{P}_k] = i\hbar$.

Let us now apply the same strategy, starting with $[\hat{Q}_k, \hat{P}_k]$. Namely, $[\hat{Q}_k, \hat{P}_k]$ is expressed in terms of expansions of its contents in n-space. As before, we shall begin by trying the "opposite" sign in the exponent, just to see what happens, *i.e.*, $\hat{p}_n e^{-ikan}$ is used instead of $\hat{p}_n e^{ikan}$:

$$\left[\hat{Q}_k, \hat{P}_k\right] = \frac{1}{N}\sum_{n,n'}\left[\hat{q}_n e^{-ikan}, \hat{p}_{n'} e^{-ikan'}\right] \tag{1.25}$$

$$= \frac{1}{N}\sum_{n,n'}\left[\hat{q}_n, \hat{p}_{n'}\right] e^{-ika(n+n')} \ . \tag{1.26}$$

Using $[\hat{q}_n, \hat{p}_{n'}] = i\hbar\delta_{n,n'}$ enables this to be written

$$\left[\hat{Q}_k, \hat{P}_k\right] = i\hbar\frac{1}{N}\sum_{n,n'}\delta_{n,n'} e^{-ika(n+n')} \tag{1.27}$$

$$= i\hbar\frac{1}{N}\sum_{n} e^{-i2kan} \ . \tag{1.28}$$

The sum is equal to zero, resulting in a dramatic inconsistency because the left hand side is nonzero. Again, had the phase progression given by eqn (1.21) been used, everything would have fallen into place, yielding $[\hat{Q}_k, \hat{P}_k] = i\hbar$.

The bottom line is that the alternate phase choice does not give correct physics, whereas the one given in eqns (1.20) and (1.21) works. Thus, we have arrived at the Fourier transformations between n-space and k-space. The transform pairs q_n / Q_k and p_n / P_k, and important properties of the adjoints of $\hat{Q}_k$ and $\hat{P}_k$, are listed in the box below [eqns (1.33) – (1.35)]. Though quantum mechanics has been used to establish the phase progression, this result can also be obtained using classical physics. After all, the phase choice of eqns (1.20) and (1.21) is perfectly acceptable, and correct results are obtained without actually ever *applying* quantum mechanics.

Classical Route

To establish the correct phase progression by using classical mechanics, the Poisson bracket of q_n and p_n is examined. Poisson brackets will be denoted using squiggly brackets. Note that this is unrelated to the use of squiggly brackets to denote anticommutators. The Poisson bracket of two functions of q_n and p_n [*i.e.*, $f(q_n, p_n)$ and $g(q_n, p_n)$] is defined as

$$\{f, g\} \equiv \frac{\partial f}{\partial q_n}\frac{\partial g}{\partial p_n} - \frac{\partial f}{\partial p_n}\frac{\partial g}{\partial q_n} \ . \tag{1.29}$$

It follows immediately that $\{q_n, p_n\} = 1$, whereas $\{q_n, p_{n'}\} = 0$ for $n \neq n'$ on physical grounds, and mathematically because $\partial p_{n'} / \partial p_n = 0$. The Poisson bracket of q_n and p_n is now expressed in terms of Fourier expansions. As in the quantum mechanical case, we shall start by trying $e^{ikan} P_k$, just to see what happens:

$$\{q_n, p_n\} = \frac{1}{N} \sum_{k,k'} \{Q_k e^{ikan}, P_{k'} e^{ik'an}\}. \tag{1.30}$$

Next, use the fact that $k = k'$ needs to be satisfied in order to achieve a nonzero value of the commutator on the right hand side. This yields

$$\{q_n, p_n\} = \frac{1}{N} \sum_{k} \{Q_k, P_k\} e^{i2kan} \tag{1.31}$$

$$= \frac{1}{N} \{Q_k, P_k\} \underbrace{\sum_{k} e^{i2kan}}_{=0} \tag{1.32}$$

Again, an inconsistency brought about by the choice of phase progression is revealed. Of course the problem is remedied straightaway with the phase progression given in eqns (1.20) and (1.21). Most importantly, classical physics yields the correct phase progression. Useful expressions are given in the box below. Everything is classical, as there is no need at the present to switch to quantum mechanics.

$$q_n = \frac{1}{\sqrt{mN}} \sum_{k} Q_k e^{ikan} \qquad p_n = \sqrt{\frac{m}{N}} \sum_{k} P_k e^{-ikan} \tag{1.33}$$

$$Q_k = \sqrt{\frac{m}{N}} \sum_{n} q_n e^{-ikan} \qquad P_k = \frac{1}{\sqrt{mN}} \sum_{n} p_n e^{ikan} \tag{1.34}$$

$$Q_k^{\dagger} = Q_{-k} \qquad P_k^{\dagger} = P_{-k} \tag{1.35}$$

Strictly speaking, $Q_k^* = Q_{-k}$ and $P_k^* = P_{-k}$ should be used, because the switch from classical to quantum mechanics has not yet been carried out. The use of the adjoint symbol † is in anticipation of this switch.

Non-Hermitian Operators Q_k and P_k

When quantum mechanics is switched on, the displacement q_n and its canonically conjugate momentum p_n, being observables, become Hermitian operators. We have seen that q_n and p_n can each be expressed as an expansion over all allowed k values, with the phase kan being the discrete counterpart to kx, where k and x vary continuously. These expansions are referred to as discrete Fourier series, or discrete Fourier transforms.[13] We will see that the quantum mechanical case results in vibrational wave functions at each site. For example, imagine a ground vibrational state wave function centered on each yellow dot,[14] with the only site-to-site difference between wave functions being a phase progression of ka in going from one site to the next.

Referring to eqns (1.33) – (1.35), Q_k and P_k are Fourier expansion coefficients. As discussed earlier, there is no reason to expect them to become Hermitian operators when q_n and p_n become Hermitian operators. Indeed, they do not. There is another big difference between the pairs q_n / p_n and Q_k / P_k. The former are local to a site, whereas Q_k and P_k extend over the entire lattice. They are discrete fields that do not correspond to observables. It was noted earlier, and will be seen in more detail below, that they satisfy essentially the same commutator relation as do q_n and p_n. However, their spatial extent is enormous compared to that of a q_n / p_n pair.

All of this might sound harmless, and perhaps it is, but at the same time it is subtle. Again, quantum mechanically, for a specific k value, Q_k and P_k extend over the entire lattice. On the other hand, if all k values are used, vibrational excitation can be created at a specific site. Between these extremes, if vibrational excitation is spread over a number of sites, a less broad distribution of k values is used. In this sense, there are no surprises. Advantage emerges when the H given by eqn (1.3) is transformed to k-space. When this is done, a compact form is obtained in which H is interpreted as a collection of independent oscillators, one for each k value.

---

[13] This mathematics is used in algorithms that carry out the Fast Fourier Transform (FFT), which is in widespread use nowadays.

[14] This scenario presupposes that Ω_0 is sufficiently larger than Ω to localize the mass points near the equilibrium sites.

1.3. Turning on Quantum Mechanics

The *k*-space counterparts of q_n and p_n, Q_k and P_k, are referred to as normal coordinate and normal momentum, respectively. Together they are used to represent the normal modes of the lattice. Each *k* value has associated with it one normal mode. You are familiar with the normal modes of polyatomic molecules. Each of these normal modes consists of atomic displacements from their respective equilibrium positions, with each atom oscillating at the normal mode frequency. Any one atomic displacement corresponds to a quantum mechanical observable. However, the collection of displacements associated with a normal mode does not correspond to a quantum mechanical observable. In other words, normal mode energies are observables, but not the multitudes of atomic displacements and their respective phases that constitute the normal modes.

The same idea applies here. In the case of a crystal, a given *k* value defines a normal mode whose spatial extent spans the entire lattice. No single measurement can determine all of the atomic displacements at one time. Note in particular the relations $Q_k{}^\dagger = Q_{-k}$ and $P_k{}^\dagger = P_{-k}$. You are used to thinking of coordinate and momentum as self-adjoint (Hermitian) operators. Indeed, q_n and p_n will become Hermitian operators when quantum mechanics is switched on by assigning operator status to them in the commutator relation $\left[\hat{q}_n, \hat{p}_{n'}\right] = i\hbar\delta_{n,n'}$. However, as explained earlier, their *k*-space counterparts, Q_k and P_k, will be non-Hermitian.

Fourier analysis has a long and remarkably successful history of facilitating problem solving by enabling one to work in the more expedient "space" for a given task. In the present case, we know that the system has oscillation frequencies, and that these frequencies are harmonic. Two frequencies Ω_0 and Ω appear in the Hamiltonian given by eqn (1.3), and it must be possible to represent the excitations in terms of harmonic oscillators with one frequency for each value of $|k|$. This lies at the heart of transforming the problem to *k*-space.

It is not possible to combine two harmonic vibrations in a way that results in anharmonicity. Consequently, we anticipate harmonic oscillators whose frequencies depend on *k* amd somehow contain Ω_0 and Ω. Also, from the symmetry of the Born and von Kármán loop, we see that there cannot be different frequencies for two *k* values having the same magnitude but opposite sign.

The essence of normal mode analysis is that frequencies are obtained for the different symmetry groups of a system. Here we have the symmetry of the loop with its twofold degeneracy, so *k* values that differ only in sign will appear together with a common frequency. Recall from Part IV.A, and from introductory molecular orbital and group theory, that these doubly degenerate pairs belong to respective symmetry groups. This ensures that each value of $|k|$ has its own frequency. The bottom line is that we know *a priori* that the Hamiltonian, when represented in *k*-space, will be a sum of harmonic oscillators, with each value of $|k|$ having its own frequency.

So far the development has been based entirely on classical physics. Quantum mechanics has not entered in the least. It was enlisted as one way to establish a phase progression, but even then it turned out that classical physics performed the task just as well. The

transition to quantum mechanics is now made through the introduction of the usual commutation relations:

$$\left[\hat{q}_n\,,\,\hat{p}_{n'}\right] = i\delta_{n,n'} \qquad \left[\hat{q}_n,\hat{q}_{n'}\right] = 0 = \left[\hat{p}_n,\hat{p}_{n'}\right], \tag{1.36}$$

where the convention $\hbar = 1$ is used.

It is not essential to introduce quantum mechanics at this point. It could be done just before introducing second quantization. It is noteworthy that a great deal of the theory is classical mechanics and Fourier analysis. It is not difficult to use eqn (1.36) to work out the corresponding commutation relation for $\hat{Q}_k$ and $\hat{P}_{k'}$. This is carried out in the box below. Recall that a similar mathematical maneuver was used to establish the correct phase relations between Q_k and P_k.

$$[\hat{Q}_k\,,\,\hat{P}_{k'}] = \frac{1}{N}\left[\sum_n e^{-ikan}\hat{q}_n\,,\,\sum_{n'} e^{ik'an'}\hat{p}_{n'}\right] = \frac{1}{N}\sum_{n,n'} e^{-ikan}e^{ik'an'}[\hat{q}_n\,,\,\hat{p}_{n'}]$$

$[\hat{Q}_k\,,\,\hat{P}_{k'}]$ vanishes unless $k = k'$ and $n = n'$. The latter follows from $[\hat{q}_n\,,\,\hat{p}_{n'}] = i\delta_{n,n'}$.

The sum over n and n' reduces to a sum over n: $[\hat{Q}_k\,,\,\hat{P}_{k'}] = \frac{i}{N}\sum_n e^{-i(k-k')an} = i\delta_{k,k'}$.

The commutators $[\hat{Q}_k\,,\,\hat{Q}_{k'}]$ and $[\hat{P}_k\,,\,\hat{P}_{k'}]$ follow along similar lines.

The above results are summarized as

$$[\hat{Q}_k\,,\,\hat{P}_{k'}] = i\delta_{k,k'} \qquad [\hat{Q}_k\,,\,\hat{Q}_{k'}] = 0 = [\hat{P}_k\,,\,\hat{P}_{k'}]\,. \tag{1.37}$$

The commutation relations for the normal coordinate and momentum, $\hat{Q}_k$ and $\hat{P}_k$, have a similar appearance to those for $\hat{q}_n$ and $\hat{p}_n$. In going from the site-specific $\hat{q}_n$ and $\hat{p}_n$ to the k-specific $\hat{Q}_k$ and $\hat{P}_k$ the n and n' indices are replaced by k and k'. The distinction is important. The n and n' indices refer to *specific lattice sites*, whereas each k value subsumes *all lattice sites*. For a simple analogy, think of the asymmetric stretch mode of CO_2. The nuclei move collectively at the normal mode frequency. The same idea applies here. The normal coordinate and momentum subsume the collective dynamics of all particles.[1]

Now that we have obtained the expansion of $\hat{q}_n$ and $\hat{p}_n$ in terms of $\hat{Q}_k$ and $\hat{P}_k$, the Hamiltonian given by eqn (1.3) is transformed to the $\hat{Q}_k$ / $\hat{P}_k$ (k-space) representation by introducing the expansions given in eqn (1.33). This leads to a compact form for the

[1] The way the model is constructed, one of the normal modes will have a frequency whose value is zero. For a linear 1D chain this corresponds to overall translation of the chain as a whole. For the loop, there is simply no angular momentum ($J = 0$). However, keep in mind that the loop is a mathematical construct.

Hamiltonian represented in *k*-space. As mentioned above, this maneuver has nothing to do with quantum mechanics. We could just as easily switch on quantum mechanics after deriving the *k*-space Hamiltonian.

Converting the Hamiltonian given by eqn (1.3) to its *k*-space counterpart will reveal the spectrum of lattice vibrational frequencies (normal modes), ω_k. Again, this is not much different conceptually than finding the normal modes and frequencies of a polyatomic molecule. In the present case, however, the task is much easier because of the discrete translational symmetry of the 1D lattice. Imagine the amount of work that would be required to obtain the normal modes of a large complicated polyatomic molecule.

I have nothing against mathematics or attention to detail, but getting bogged down repeatedly disrupts one's train of thought. Consequently, what I consider tedious math will hereafter be placed in boxes. This practice will streamline the presentation in a manner that hopefully lets the main ideas come across, while at the same time providing the intermediate steps, so results do not appear out of thin air. In keeping with this strategy, details of the conversion of eqn (1.3) to its *k*-space counterpart are given in the box below.

k-Space Representation of H

Eqn (1.3): $\hat{H} = \frac{1}{2}\sum_{n=1}^{N}\left\{\frac{{p_n}^2}{m} + m\Omega_0{}^2 {q_n}^2 + m\Omega^2\left(q_n - q_{n-1}\right)^2\right\}$. Using eqn (1.33), $\hat{H}$ becomes

$$\frac{1}{2N}\sum_n\left\{\sum_{k,k'}\hat{P}_k e^{-ikan}\hat{P}_{k'}e^{-ik'an} + \Omega_0^2\sum_{k,k'}\hat{Q}_k e^{ikan}\hat{Q}_{k'}e^{ik'an} + \Omega^2\left(\sum_k \hat{Q}_k e^{ikan} - \sum_{k'}\hat{Q}_{k'}e^{ik'a(n-1)}\right)^2\right\}$$

$$\frac{1}{2N}\sum_{k,k'}\hat{P}_k\hat{P}_{k'}\underbrace{\sum_n e^{-i(k+k')an}}_{N\delta_{k,-k'}} \rightarrow \underbrace{\frac{1}{2}\sum_{k,k'}\hat{P}_k\hat{P}_{k'}\delta_{k,-k'}}_{\frac{1}{2}\sum_k \hat{P}_k\hat{P}_k^\dagger}}$$

$$\frac{\Omega_0^2}{2N}\sum_{k,k'}\hat{Q}_k\hat{Q}_{k'}\underbrace{\sum_n e^{i(k+k')an}}_{N\delta_{k,-k'}} \rightarrow \underbrace{\frac{\Omega_0^2}{2}\sum_{k,k'}\hat{Q}_k\hat{Q}_{k'}\delta_{k,-k'}}_{\frac{1}{2}\Omega_0{}^2\sum_k\hat{Q}_k\hat{Q}_k^\dagger}}$$

These 4 terms are given by:

$$\frac{1}{2}\Omega^2\sum_{k,k'}\hat{Q}_k\hat{Q}_{k'}\delta_{k,-k'}\left(2 - e^{ika} - e^{-ika}\right)$$

$$= \frac{1}{2}\Omega^2\sum_k \hat{Q}_k\hat{Q}_k^\dagger \underbrace{2\left(1-\cos ka\right)}_{2\sin^2(ka/2)}$$

$$\frac{1}{2}4\Omega^2\sum_k\hat{Q}_k\hat{Q}_k^\dagger\sin^2(ka/2)$$

$$\frac{1}{2}\sum_k\hat{Q}_k\hat{Q}_k^\dagger\underbrace{\left(\Omega_0{}^2 + 4\Omega^2\sin^2(ka/2)\right)}_{\omega_k{}^2}$$

You should go through this exercise step by step. It will not kill you, and upon completion you will appreciate further the logic that underlies the transition to *k*-space. The important results are given below in eqns (1.38) and (1.39).

$$\hat{H} = \tfrac{1}{2}\sum_k \left(\hat{P}_k \hat{P}_k^{\dagger} + \omega_k^2 \hat{Q}_k \hat{Q}_k^{\dagger} \right) . \tag{1.38}$$

The dispersion relation is the relationship between ω_k and k. From the lower right of the large box on the previous page, we see that it is given by

$$\omega_k^2 = \Omega_0^2 + 4\Omega^2 \sin^2(ka/2) . \tag{1.39}$$

The reason eqn (1.39) is referred to as a dispersion relation comes from wave propagation. The phase and group velocities of waves, ω / k and $\partial\omega / \partial k$, respectively, are obtained from this expression, and these tell us things like how fast a wave spreads as it propagates and how fast energy is transported (Appendix 2). A point to keep in mind – and one that often goes without mention – is that group velocity only has physical meaning for localized wave packets. One can evaluate $\partial\omega / \partial k$ at a specific k value, but a single k value is spatially delocalized, so energy transport is not defined. A wave packet provides a picture of what really happens. Should the wave packet get crazy (for example bifurcate into two packets) then group velocity loses meaning, but for a sound physical reason.

As mentioned a number of times, the representation of the Hamiltonian given by eqn (1.38) provides great advantage over the one given by eqn (1.3). To solve eqn (1.3) directly, one must deal with the *N* individual mass points, each of which is coupled to its nearest neighbors and to the laboratory reference frame. Consequently, the masses are all coupled together. On the other hand, with eqn (1.38) the only possible couplings are between k and $-k$ (recall that $Q_k^{\dagger} = Q_{-k}$ and $P_k^{\dagger} = P_{-k}$).[2]

The summation in eqn (1.38) is over positive and negative k, and it is tempting to rewrite this in a manner such that k assumes only positive values. For the time being, the degenerate $\pm |k|$ levels shall be retained. They correspond to phase progression in opposite senses on the ring shown in Fig. 3. This phase progression does not correspond to conserved (angular) momentum because the symmetry is discrete rather than continuous. In Part IV.A, the particle-on-a-ring with a periodic potential $V(\phi)$ was examined assiduously. Recall that it was shown that when $V(\phi)$ is strongly binding, there is a quite small circulating flux, whereas when $V(\phi)$ is weak, the motion on the ring is almost unhindered.

---

[2] Actually, the model is simple enough to yield an analytical solution when solved in coordinate space if Ω_0 is set equal to zero. However, with $\Omega_0 \neq 0$ and including other couplings, such as between more distant sites, as well as anharmonicity, it would be, for all practical purposes, impossible to obtain an analytical solution. Matrix diagonalization would be required. Working in *k*-space, however, it would be straightforward to derive an expression for $\omega_k(k)$ (the dispersion relation) and represent the Hamiltonian as a sum over normal modes.

1.4. Dispersion Relation

With the above Hamiltonian and the dispersion relation plotted in Fig. 8, we can begin to see our way to the eigenvalues. Namely, for each lattice normal mode (k value), there is a frequency ω_k, with the understanding that the lattice normal mode with $-k$ also has this eigenvalue. It is, of course, understood that we are taking the positive square root of $\omega_k{}^2 = \Omega_0{}^2 + 4\Omega^2 \sin^2(ka/2)$ given by eqn (1.39). The upper and lower curves in the figure ($\Omega_0 = 3\Omega$ and $\Omega_0 = 0$) are referred to as the optical and acoustic branches, respectively. There is an ω_k region that is not allowed. It is indicated with a thick vertical arrow in Fig. 8. The crystal cannot undergo oscillation at frequencies that lie in this gap, as no normal modes have these frequencies. Likewise, there are no normal modes with $\omega_k > \sqrt{13}\,\Omega \approx 3.6\,\Omega$, as seen in Fig. 8.

In our toy model, the acoustic and optical branches are obtained by assigning values to Ω_0. Namely, $\Omega_0 = 0$ gives the acoustic branch, and $\Omega_0 \neq 0$ gives the optical branch. This differs from lattice models introduced later, but nonetheless displays the qualitative character of the acoustic and optical branches

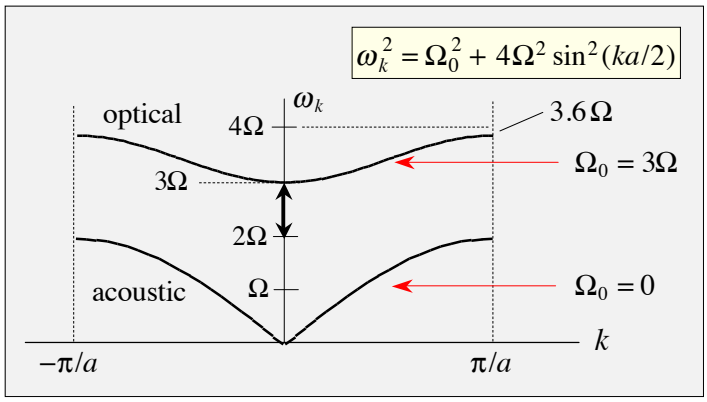


Figure 8. The dispersion relation given by eqn (1.39) is shown for: $\Omega_0 = 0$ and $\Omega_0 = 3\Omega$, *i.e.*, acoustic and optical branches, respectively. The first Brillouin zone lies between the vertical dashed lines. It is interesting to compare these curves to those of the electronic band structure of a 1D lattice. The $\Omega_0 = 0$ curve is to scale; the $\Omega_0 = 3\Omega$ curve is an educated guess.

Optical phonons have what can be described as *local-mode character*, in the sense that vibrations are about the equilibrium positions of the groups of atoms that constitute a repeating element.[3] Think of a CO_2 crystal. The high frequency intramolecular CO_2 vibrations are local modes. They give rise to optical phonons. Of course, a CO_2 crystal is a 3D object, so there are more modes than those indicated in Fig. 8. As discussed later, for CO_2 there are 3 acoustic modes and 6 optical modes. You can think of these as correlating with a free CO_2 molecule's 3 translational degrees of freedom and 6 vibration-rotation degrees of freedom.

[3] Optical phonons got their name because of how these modes interact with radiation in an ionic crystal. The name stuck, and it is now used with all crystal types.

Notice that the group velocity, $v_g = \partial\omega / \partial k$, is lower for the optical than the acoustic branch except at the edge of the Brillouin zone, where each vanishes. Acoustic phonons, as the name implies, are sound waves. In general, optical branch wave packets carry local oscillation through the lattice. It is not easy for high frequency vibrations to move along the lattice because they must hop between sites. This is why group velocity is smaller for optical branches.

Referring to Fig. 8, a number of limiting cases are intuitive. When $\Omega = 0$, communication between masses at adjacent sites is gone. In this case, the problem reduces to the trivial case of N independent oscillators. Relative phase between displacements at adjacent sites becomes irrelevant in the sense that a given displacement is not influenced by the phases of its neighbors. The dispersion curve in Fig. 8 can be tossed out, and we are left with $\omega_k = \Omega_0$. Alternatively, when $\Omega_0 = 0$, the system behaves like an elastic rod for smaller k values (a grainy one because of the discrete mass points), as binding to the equilibrium positions (yellow dots in Fig. 1) is gone. As mentioned earlier, with $\Omega_0 = 0$, we have the simple example of a monatomic lattice with equal spring constants. The *elastic limit* is obtained by converting the mass density of the discrete lattice (m / a) to a continuous mass distribution with the same density. In this case the dispersion relation is linear: $\omega_k = v_0 k$, where v_0 is the speed of sound. The edge of the Brillouin zone approaches $k = \infty$ as $a \to 0$.

Our toy model represents both local and inter-site vibrations, but so do other lattice models. When an ionic crystal such as NaCl is modeled (Section 7), it will be seen that the optical branch looks different than the one shown in Fig. 8. Acoustic branches always look pretty much like the one shown in Fig. 8. The main difference is scaling the vertical axis. Optical branches are relatively flat and their shapes can vary dramatically depending on the parameters of a given model, for example, from a maximum at $k = 0$ to a minimum at $k = 0$. Notice that for the optical branches indicated in Fig. 9 the group velocity $\partial\omega / \partial k$ is negative throughout the region $0 \le k \le \pi / a$.

Later our model will be extended to 3D. Currently, it is 1D, so waves, displacements, and momenta are all longitudinal. A transverse coordinate does not exist. Were the system 3D, a given propagation direction would accommodate transverse as well as longitudinal phonons. This is indicated schematically in Fig. 9, which is for 2 atoms per unit cell. When there are more than 2 atoms per unit cell, say g atoms per unit cell, there are still 3 acoustic branches but there are $3g - 3$ optical branches. If the repeating unit is

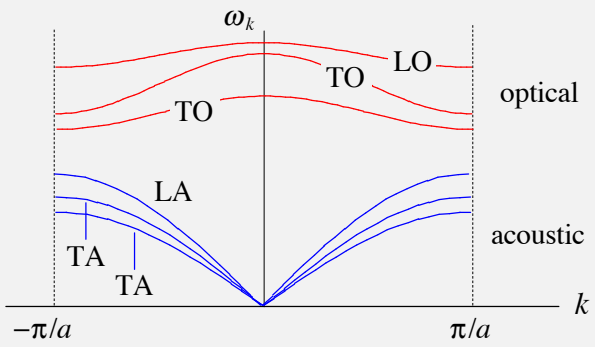


Figure 9. When a 3D crystal has two atoms per unit cell, there are 3 acoustic and 3 optical branches. Of the acoustic branches, 2 are transverse and 1 is longitudinal (TA and LA, respectively). Likewise, there are 2 transverse and 1 longitudinal optical branches (TO and LO, respectively).

polyatomic and intermolecular coupling is weak, the polyatomic vibrations manifest as optical branches, whereas the acoustic branches are determined by the coupling between adjacent polyatomic species.

There are two transverse directions for a given propagation direction. Consequently, optical and acoustic branches each contain twice as many transverse as longitudinal branches. The acronyms are intuitive: LO and TO for longitudinal and transverse optical, and LA and TA for longitudinal and transverse acoustic. Let us now find the complete spectrum of eigenvalues by solving the Schrödinger equation using the Hamiltonian given by eqn (1.38).

1.5. Second Quantization

It is assumed that you are familiar with raising and lowering operators, for example, as used in angular momentum theory and the quantum mechanical harmonic oscillator. If not, go to the end of this section where a description of how this works is given. In addition, you should consult a text that discusses this in more detail. Equation (1.38) shows that we have a harmonic oscillator here as well. This one is in terms of the normal coordinates and momenta $\hat{Q}_k$ and $\hat{P}_k$. The only mathematical difference between this harmonic oscillator and the one that uses $\hat{q}_n$ and $\hat{p}_n$ is that here we have to deal more carefully with the adjoints $\hat{Q}_k{}^\dagger$ and $\hat{P}_k{}^\dagger$, because $\hat{Q}_k$ and $\hat{P}_k$ are not Hermitian. Specifically, $\hat{Q}_k{}^\dagger = \hat{Q}_{-k}$, which is certainly not $\hat{Q}_k$, and likewise for $\hat{P}_k{}^\dagger$.

This is interesting. What consequences might there be from non-Hermitian $\hat{Q}_k$ and $\hat{P}_k$? We will have the answer soon. For the time being, the essential equations will be developed. This will enable us to move quickly to the result, which will then be discussed. To achieve this, annihilation and creation operators are introduced. Their commutation relations are obtained with a minimum amount of algebra. This is left as an exercise. The relevant formulas are

$$\hat{a}_k = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k \hat{Q}_k + i\hat{P}_k{}^\dagger\right) \qquad \hat{a}_k{}^\dagger = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k \hat{Q}_k{}^\dagger - i\hat{P}_k\right) \tag{1.40}$$

$$\hat{Q}_k = \frac{1}{\sqrt{2\omega_k}}\left(\hat{a}_k + \hat{a}_{-k}{}^\dagger\right) \qquad \hat{P}_k = -i\sqrt{\frac{\omega_k}{2}}\left(\hat{a}_{-k} - \hat{a}_k{}^\dagger\right) \tag{1.41}$$

$$\left[\hat{a}_k, \hat{a}_k{}^\dagger\right] = 1 .$$

Note that use has been made of the relations listed in eqn (35): $\hat{Q}_k{}^\dagger = \hat{Q}_{-k}$ and $\hat{P}_k{}^\dagger = \hat{P}_{-k}$. Next, the above expressions for $\hat{Q}_k$ and $\hat{P}_k$ in terms of $\hat{a}_k$ and $\hat{a}_k{}^\dagger$ are used with the Hamiltonian given by eqn (1.38). Following the algebra worked out in the box below, the Hamiltonian emerges as the compact expression given by eqn (1.42).

$$\hat{H} = \tfrac{1}{2}\sum_k \left(\hat{P}_k\hat{P}_k{}^\dagger + \omega_k{}^2\,\hat{Q}_k\hat{Q}_k{}^\dagger\right) = \sum_k \tfrac{\omega_k}{4}\left\{\left(\hat{a}_{-k} - \hat{a}_k{}^\dagger\right)\left(\hat{a}_{-k}{}^\dagger - \hat{a}_k\right) + \left(\hat{a}_k + \hat{a}_{-k}{}^\dagger\right)\left(\hat{a}_k{}^\dagger + \hat{a}_{-k}\right)\right\}$$

$$= \sum_k \tfrac{\omega_k}{4}\left\{\hat{a}_{-k}\hat{a}_{-k}^\dagger - \cancel{\hat{a}_{-k}\hat{a}_k} - \cancel{\hat{a}_k^\dagger\hat{a}_{-k}^\dagger} + \hat{a}_k^\dagger\hat{a}_k + \hat{a}_k\hat{a}_k^\dagger + \cancel{\hat{a}_k\hat{a}_{-k}} + \cancel{\hat{a}_{-k}^\dagger\hat{a}_k^\dagger} + \hat{a}_{-k}^\dagger\hat{a}_{-k}\right\}$$

Using $\hat{a}_k\hat{a}_k^\dagger = 1 + \hat{a}_k^\dagger\hat{a}_k$, and $\hat{a}_{-k}\hat{a}_{-k}^\dagger = 1 + \hat{a}_{-k}^\dagger\hat{a}_{-k}$: $H = \sum_k \frac{\omega_k}{2}\left(\hat{a}_k^\dagger\hat{a}_k + \hat{a}_{-k}^\dagger\hat{a}_{-k} + 1\right)$

The above summation is over positive and negative k values. Converting this summation to one over only positive k values yields a neat form for the Hamiltonian and corresponding energies

$$\hat{H} = \sum_k \omega_k\left(\hat{a}_k{}^\dagger\,\hat{a}_k + \tfrac{1}{2}\right) \tag{1.42}$$

$$E = \sum_k \omega_k\left(n_k + \tfrac{1}{2}\right). \tag{1.43}$$

The creation and annihilation operators: $\hat{a}_k{}^\dagger$ and $\hat{a}_k$, respectively, have well-known properties that you can find in any number of texts. A review of raising and lowering operators for the harmonic oscillator is given for your convenience in the next two pages, so you can refer to these results without searching through texts.

Exercise: Show how eqn (1.42), which sums only over positive k, follows from the last line of the large box, where the summation is carried out over all positive and negative k values.

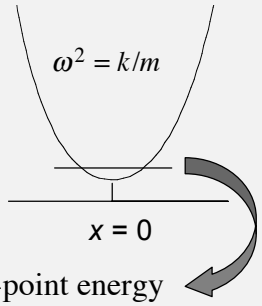
Harmonic Oscillator: Raising and Lowering Operators

This is a concise summary of the mathematics that underlies the use of raising and lowering (also called creation and annihilation) operators for the harmonic oscillator. Start with the harmonic oscillator Hamiltonian:

$$H=\frac{kx^2}{2}+\frac{p^2}{2m}=\omega\left(\frac{m\omega}{2}x^2+\frac{p^2}{2m\omega}\right)=\hbar\omega\left\{\left(\sqrt{\frac{m\omega}{2\hbar}}x\right)^2+\left(\frac{p}{\sqrt{2m\omega\hbar}}\right)^2\right\}$$

Factoring this and introducing the operators $\hat{a}^\dagger$ and $\hat{a}$ gives

$$\hbar\omega\underbrace{\left(\sqrt{\frac{m\omega}{2\hbar}}x-ip\frac{1}{\sqrt{2m\omega\hbar}}\right)}_{\hat{a}^\dagger}\underbrace{\left(\sqrt{\frac{m\omega}{2\hbar}}x+ip\frac{1}{\sqrt{2m\omega\hbar}}\right)}_{\hat{a}}\underbrace{-i\frac{\omega}{2}[x,p]}_{\frac{1}{2}\hbar\omega}$$



$$H=\hbar\omega\left(\hat{a}^\dagger\hat{a}+\tfrac{1}{2}\right)$$

$\hat{a}^\dagger\hat{a}$ = number operator

zero-point energy

It is easy to show that $[\hat{a},\hat{a}^\dagger]=1$ and that $[H,\hat{a}]=-\hbar\omega\hat{a}$ (see white box).

Using these results we write: $[H,\hat{a}]\psi=-\hbar\omega\hat{a}\psi=H\hat{a}\psi-\hat{a}H\psi=H\hat{a}\psi-E\hat{a}\psi$

Thus, $H(\hat{a}\psi)=(E-\hbar\omega)(\hat{a}\psi)$.

Given that E is the eigenvalue of ψ, we see that $E-\hbar\omega$ is the eigenvalue of $\hat{a}\psi$.

Therefore $\hat{a}\psi$ is proportional to an eigenfunction with eigenvalue $E-\hbar\omega$.

Likewise for $\hat{a}^\dagger$ and $E+\hbar\omega$.

Eigenvalues are separated by $\hbar\omega \Rightarrow E=\hbar\omega\left(v+\tfrac{1}{2}\right)$.

$$[a,a^\dagger]=\left(\sqrt{\frac{m\omega}{2\hbar}}x+ip\frac{1}{\sqrt{2m\omega\hbar}}\right)\left(\sqrt{\frac{m\omega}{2\hbar}}x-ip\frac{1}{\sqrt{2m\omega\hbar}}\right)$$
$$-\left(\sqrt{\frac{m\omega}{2\hbar}}x-ip\frac{1}{\sqrt{2m\omega\hbar}}\right)\left(\sqrt{\frac{m\omega}{2\hbar}}x+ip\frac{1}{\sqrt{2m\omega\hbar}}\right)$$
$$=-\frac{i}{\hbar}[x,p]=1$$

$$[H,\hat{a}]=\hbar\omega\left[\hat{a}^\dagger\hat{a},\hat{a}\right]=\hbar\omega(\hat{a}^\dagger\hat{a}-\hat{a}\hat{a}^\dagger)\hat{a}=-\hbar\omega\hat{a}$$

Next, the exact expressions for $\hat{a}^\dagger|v\rangle$ and $\hat{a}|v\rangle$ are obtained.

Start with $\hat{a}^\dagger|v\rangle = c_+|v+1\rangle$ and $\langle v|\hat{a} = c_+{}^*\langle v+1|$ and take the inner product:

$\langle v|\hat{a}\hat{a}^\dagger|v\rangle = \langle v|\hat{a}^\dagger\hat{a}+1|v\rangle = (v+1)\langle v|v\rangle = c_+{}^* c_+\langle v+1|v+1\rangle = |c_+|^2$

Choose $c_+ = \sqrt{v+1}$. The analogous calculation for c_- and $c_-{}^*$ gives $c_- = \sqrt{v}$.

Thus: $\hat{a}^\dagger|v\rangle = \sqrt{v+1}\,|v+1\rangle$ and $\hat{a}|v\rangle = \sqrt{v}\,|v-1\rangle$.

To obtain the harmonic oscillator wave functions, apply the lowering operator to the ground state to annihilate it:

$$\hat{a}\psi = 0 \Rightarrow -\frac{m\omega}{\hbar}x\psi = \frac{\partial\psi}{\partial x} \Rightarrow \int\frac{d\psi}{\psi} = -\frac{m\omega}{\hbar}\int dx\, x = -\frac{m\omega}{2\hbar}x^2$$

Normalizing gives the ground state wave function : $\left(\frac{m\omega}{\pi\hbar}\right)^{1/4}\exp\left(-\frac{m\omega}{2\hbar}x^2\right)$.

Successive application of $a^\dagger$ (plus normalization) yields the entire spectrum of harmonic oscillator wave functions.

The above results are also applicable to the lattice modes:

$\hat{a}_k^\dagger|n_k\rangle = \sqrt{n_k+1}\,|n_k+1\rangle$

$\hat{a}_k|n_k\rangle = \sqrt{n_k}\,|n_k-1\rangle$

$\hat{a}_k^\dagger\hat{a}_k|n_k\rangle = n_k|n_k\rangle$

$\hat{a}_k^\dagger\hat{a}_k^\dagger\hat{a}_k^\dagger|0\rangle = \sqrt{3}\sqrt{2}\sqrt{1}\,|3\rangle \Rightarrow |n_k\rangle = \frac{(\hat{a}_k^\dagger)^{n_k}|0\rangle}{\sqrt{n_k!}}$

Generalizing to any number of modes:

$$|n_{k_1}, n_{k_2}, \ldots\rangle = \frac{\left((\hat{a}_{k_1}^\dagger)^{n_{k_1}}(\hat{a}_{k_2}^\dagger)^{n_{k_2}}\ldots\right)|0,0,\ldots\rangle}{\sqrt{(n_{k_1}!)(n_{k_2}!)\ldots}}$$

1.6. Interpretation

We are now in a position to interpret the results. The importance of this cannot be overstated. You may wish to review the closely related problem of the 1D harmonic oscillator before proceeding, for example, by revisiting the summary at the end of the previous section. Detailed explanations can be found in a number of texts.

The Hamiltonian given by eqn (1.42) acts on vectors $|n_k\rangle$ in a number-valued Fock space, and in so doing creates quanta with energy ω_k for each k value. But what exactly are these vectors and how are we to interpret operations such as $\hat{a}_k{}^\dagger \hat{a}_k |n_k\rangle$ and $\hat{a}_k{}^\dagger |n_k\rangle$? To see what is going on, let us begin by expanding $\hat{a}_k{}^\dagger$ in terms of $\hat{q}_n$ and $\hat{p}_n$. The expression for $\hat{a}_k{}^\dagger$ in terms of $\hat{Q}_k{}^\dagger$ and $\hat{P}_k$ that is given by eqn (1.40) is reproduced here as eqn (1.44):

$$\hat{a}_k{}^\dagger = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k \hat{Q}_k{}^\dagger - i\hat{P}_k\right). \tag{1.44}$$

Expressing $\hat{Q}_k{}^\dagger$ and $\hat{P}_k$ as expansions over $\hat{q}_n$ and $\hat{p}_n$ [see eqn (1.34)] yields

$$\hat{a}_k{}^\dagger = \frac{1}{\sqrt{2\omega_k}}\left(\omega_k\sqrt{\frac{m}{N}}\sum_n e^{ikan}\hat{q}_n - i\frac{1}{\sqrt{mN}}\sum_n e^{ikan}\hat{p}_n\right) \tag{1.45}$$

$$= \frac{1}{\sqrt{N}}\sum_n e^{ikan}\left(\sqrt{\frac{m\omega_k}{2\hbar}}\,\hat{q}_n - i\frac{1}{\sqrt{2m\omega_k\hbar}}\,\hat{p}_n\right). \tag{1.46}$$

In going from eqn (1.44) to eqn (1.45), $\hat{q}_n{}^\dagger = \hat{q}_n$ and $(e^{-ikan})^\dagger = e^{ikan}$ have been used. In eqn (1.46), $\hbar$, which was implicit ($\hbar = 1$) until now, has been made explicit.

The term in parentheses in eqn (1.46) is familiar. It is given at the end of the previous section in the large box entitled *Harmonic Oscillator: Raising and Lowering Operators*. It is the harmonic-oscillator raising operator for the mass point associated with the n^{th} site. Note that when $\Omega_0 = 0$ the yellow dots in Fig. 1 become irrelevant, in which case n refers only to the mass points. I will continue to refer to sites, with the understanding that this implies $\Omega_0 \neq 0$.

We see that a given value of k (and therefore ω_k according to the dispersion relation) in eqn (1.46) corresponds to the same harmonic oscillator state being associated with all of the sites, with phase that progresses from site to site. The fact that the large parenthetic term in eqn (1.46) is equal to $\hat{a}_n{}^\dagger$ enables eqn (1.46) to be written

$$\hat{a}_k{}^\dagger = \frac{1}{\sqrt{N}}\sum_n e^{ikan}\hat{a}_n{}^\dagger\,. \tag{1.47}$$

Now, suppose that $\hat{a}_k^\dagger$ operates on the system. What does it do? What amplitudes does it create at the sites? How does it handle phase?

The operator $\hat{a}_k^\dagger$ singles out one value of k, but its action extends over the entire lattice. The problem is set up such that an eigenstate of the k-space Hamiltonian has equal probability density at each site, with phase progression defined by k. If the k^{th} mode contains no quanta, we label it $|0\rangle$. This is written with the understanding that $|0\rangle$ applies to the k with which we are concerned. It goes without saying that it includes all of the sites.

Figure 10 illustrates the case $\Omega_0 >> \Omega$. Mass points are bound strongly to their respective equilibrium locations, whereas inter-particle coupling is relatively weak. The harmonic oscillator ground state (shaded Gaussian) and the first two vibrationally excited states are indicated. You can appreciate that $\Omega_0 >> \Omega$ applies to Fig. 10 by noting that the wave functions shown have small overlap between sites. Recall some of the figures in Part IV.A. There, we found that a modest degree of inter-site coupling (overlap) results in a narrow energy band. Placing electrons in the single particle states (orbitals) was done in accord with the Pauli principle. In the present case, we are dealing with phonons, which are bosons. Any number of them can enter a given mode.

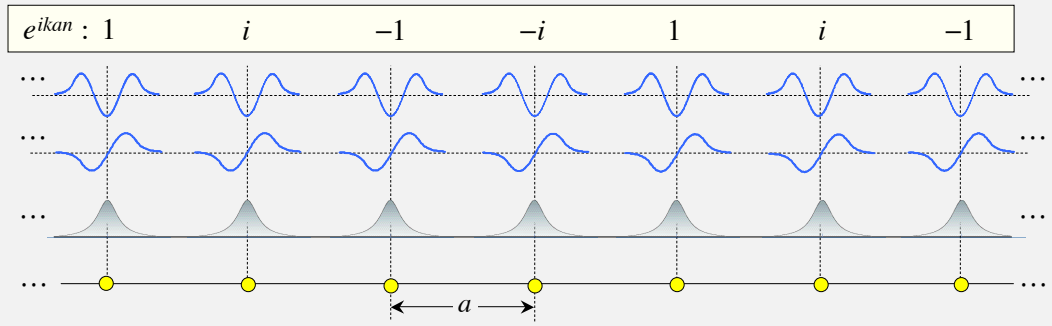
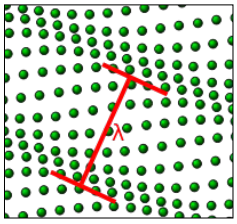


Figure 10. The $k = \pi/2a$ normal mode (say, $N = 24$ and $l = 6$) has harmonic oscillator wave functions on each site, with corresponding phases indicated at the top. Site functions are sketched (without phases) for the ground state (shaded Gaussian) and first two excited states. Nearest neighbor coupling is weak relative to site binding.

You have already dealt with something like this, but without the need for the approach used here. A diatomic molecule has equally spaced vibrational energy levels when the potential's anharmonicity is ignored. The "phonons" in this case are the vibrational quanta. They are bosons: The molecule can contain any number of these quanta.

Were we to construct a drawing in the spirit of Fig. 10, but for $\Omega_0 = 0$ (*i.e.*, for the acoustic branch), it would be quite different in appearance than Fig. 10, because in this case the particles interact with one another harmonically, but they do not interact with any underlying lattice. In one dimension the acoustic waves are, by definition, restricted to propagation in the longitudinal direction. They are compression and rarefaction waves. The pictures (below right, Google Images) use 2D ensembles of particles to illustrate compression and rarefaction waves in a lattice (left) and a fluid (right).



Simulated acoustic waves in fluids

Referring to eqn (1.47) and Fig. 10, let us consider site-to-site phase progression. For $k = \pi/2a$, which is chosen arbitrarily, phase progression is indicated with e^{ikan}. Now consider $\hat{a}_k^\dagger$ operating on $|0\rangle$. It creates harmonic oscillator levels $|1\rangle$ on each site, with phase progression $e^{ikan} = e^{in\pi/2} = i^n$. The same idea holds if $\hat{a}_k^\dagger$ operates on $|1\rangle$. It creates harmonic oscillator levels $|2\rangle$ on each site, with phase progression $e^{ikan} = e^{in\pi/2} = i^n$. In other words, phase progression is determined by k. It applies to each vibrational wave function. Thus, we have at our disposal a great piece of machinery. It enables us to create any number of quanta in any mode (k value) of our choosing.

Wave packets can be constructed by taking linear combinations of normal modes. For example, suppose we want to describe localized vibrational excitation. Imagine that a lattice made of diatomic molecules AB is photoexcited using a focused laser beam: $\mathrm{AB}(v=0)+h\nu \rightarrow \mathrm{AB}(v=1)$. It is not overly difficult experimentally to localize excitation to a diameter of ~ 30 μm. This is represented mathematically by taking enough k values to confine the $t = 0$ wave packet to this spatial region. The packet thus created then propagates. The distribution of k values will be centered on a specific value k_0 that is determined by experimental conditions. For example, if no directed momentum is imparted, k_0 will equal zero and the packet will propagate symmetrically outward from the region of excitation.

You are familiar with how the wave packet of a particle spreads as it propagates in free space. You are also familiar with the fact that an electromagnetic wave in free space propagates without changing its shape. These properties are due to the dispersion relation for free massive particles versus that of electromagnetic waves. The latter obeys $\omega = ck$, so $e^{i(kx-\omega t)} = e^{ik(x-ct)}$, so it follows that all waves travel at the speed of light regardless of their k value. On the other hand, for a massive particle in free space, we have [using $\omega = k^2/2m$ $(\hbar = 1)$]: $e^{i(kx-\omega t)} = \exp(ikx - ik^2t/2m)$. The appearance of k^2 in the argument of the exponential causes waves to disperse as they propagate. In the simplest case, this means that a packet broadens as it propagates, as described in Appendix 2.

The dispersion relation for a periodic lattice is such that there are regions where waves propagate without changing their shape, and also regions where waves change shape as they propagate. The former is identified with regions where ω is directly proportional to k that is, $\omega = vk$. In this case $e^{i(kx-\omega t)} = e^{ik(x-vt)}$, which is the same form as an electromagnetic wave except that v appears rather than the speed of light.

In this regime of propagation, the wavelength is significantly larger than the lattice spacing. You will notice from the dispersion relation in Fig. 8 that the maximum speed with which energy is transmitted occurs in the linear region of the acoustic branch.[4] Alternatively, when the wavelength becomes comparable to the lattice spacing (k moves

[4] Shock waves travel faster than the speed of sound. They are altogether another matter.

toward the edge of the Brillouin zone), the waves disperse and energy transmission slows, approaching zero at the Brillouin zone edge.

The energies of the lattice modes distinguished by the parameter k are obtained from the dispersion relation: $\omega_k{}^2 = \Omega_0{}^2 + 4\Omega^2 \sin^2(ka/2)$. For each value of k there is a frequency (energy) ω_k. Taking the square root yields: $\omega_k = \pm[\Omega_0{}^2 + 4\Omega^2 \sin^2(ka/2)]^{1/2}$. Do not worry about the $\pm$. In the classical harmonic oscillator, each is included in the solution, as the differential equation is second order in space and time. Conceptual problems would arise with negative energy solutions when solving something like the Schrödinger equation, which is first order in time. However, in our approach – in which the Hamiltonian operates in a number-valued Fock space – only the plus sign appears in the expression for the energy. This is evident in eqn (1.42), where ω_k is positive. It is also discussed at great length in Part V: *Introduction to Classical and Quantum Relativity*.[5]

There is no limit to the number of quanta that can be present in the lattice mode characterized by the parameter k. Leaving aside the zero point energy, the energy present in a given mode depends on how many phonons occupy the mode. That is, the energy for a given number of phonons (say, in the k^{th} mode) is the product of the mode energy ω_k times the occupation number n_k: $E_k = n_k\omega_k$. It is understood that the mode occupation number n_k needs to be established. For example, it might undergo thermal fluctuations, so its value does not remain constant in time. Nonetheless, we shall see that its time-averaged value can be calculated without difficulty.

The quanta are those of the number operator $a_k{}^\dagger a_k$, whose quantum mechanical properties derive from the fundamental commutation relation for particles: $[q_n, p_{n'}] = i\delta_{n,n'}$. These quanta are called *phonons* (from ancient Greek: $\varphi\omega\nu\eta$, which translates to *sound* or *voice* in English). The Russian physicist Yakov Frenkel (photo) introduced this term.

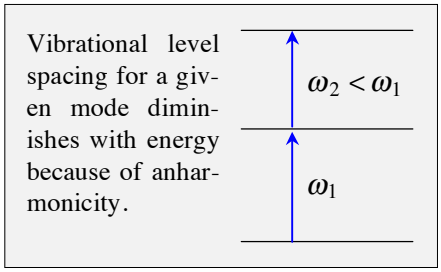
Yakov Frenkel

Lev Landau

Igor Tamm

[5] A long time ago, the sign of ω was considered problematic in relativistic quantum mechanics. Schrödinger attempted to develop a wave equation consistent with Einstein's equation for special relativity: $\omega^2 = m^2c^4 + p^2c^2$. His efforts and those of others were confounded by negative kinetic energy, negative probability density, and other egregious problems. He gave up and settled for the nonrelativistic Schrödinger equation – not bad for a consolation prize. In our model, the Hamiltonian is always positive because ω_k in eqn (1.42) is always positive.

Phonons are bosons. As mentioned earlier, as such, they have the property that there is no restriction on the number of them that can be placed in a given mode. They arise in the simple harmonic theory of the present chapter as identical quanta for a given k value. That is, leaving aside zero point energy, n_k = 0, 1, 2... corresponds to E_k = 0, ω_k, $2\omega_k$ …. Phonons are often referred to as quasiparticles. These are elementary excitations that arise when fields are quantized. They were introduced and studied assiduously by the Russian physicists Yakov Frenkel, Lev Landau (photo taken in NKVD prison following his arrest for insufficient loyalty to the Soviet State), and Igor Tamm.



Vibrational level spacing for a given mode diminishes with energy because of anharmonicity.

The Hamiltonian given by eqn (1.42) is for harmonic oscillators characterized by the parameters Ω and Ω_0 of our toy model. When anharmonic terms are included in potentials, the energy spectrum changes. However, the bosonic nature of the quanta is not affected. For example, with an anharmonic potential, adding a second phonon to a mode requires less energy than that required for the addition of the first phonon. Anharmonicity does not change the fact that there is no restriction on the number of quanta that can be present in a given mode. Because of the fact that the bosonic nature of phonons arises from the fundamental commutation relation, $[q_n, p_{n'}] = i\delta_{n,n'}$, it is unaffected by anharmonicity.[6]

Let us now discuss the form for the Hamiltonian given by eqn (1.42). When a collection of particles is described quantum mechanically, a set of single-particle Hilbert spaces can be used to keep track of both the particles and the many states that arise from their interactions with one another and with other particles. For example, in an atom, electrons interact with one another and with the nucleus. The same idea holds for molecules. In each case, the electrons and their states are described using Hilbert space.[7]

When the number of particles is variable, the number of single particle Hilbert spaces is not fixed. For example in the case of the quasiparticles that we call phonons, many of these quasiparticles can be produced in straightforward elementary events – nothing more complicated than an atom colliding with a crystal. We can write a Hilbert space for phonons, but how many of these Hilbert spaces are needed, and for each one how shall its dimension be determined? In other words, using Hilbert spaces might require some interesting juggling.

---

[6] For a molecular analog, consider a diatomic molecule. Its vibrational excitations are bosons, and the molecule can support many vibrational quanta. Though experimentally challenging, it is possible to implant a large number of quanta into a diatomic molecule by using laser excitation.

[7] The vector spaces that are used, for all practical purposes, are ray spaces, *i.e.*, composed of real kets. They single out the true states of the system, independent of things like phase factors. An example is the space of projectors, $P_n = |n\rangle\langle n|$. Multiplying $|n\rangle$ by $e^{i\varphi}$ leaves P_n unaffected. Each Hilbert vector can have phase: $|\cdot\cdot\rangle e^{i\varphi}$. Practically speaking, ray spaces function as Hilbert spaces.

A vector space is preferred that is not fixed in the number of particles or quasiparticles that can be accommodated in the theory. This unbounded vector space is called *Fock space*, after the outstanding Russian physicist and mathematician Vladimir Fock. It has a useful representation in which creation and annihilation operators act on a basis of occupation numbers.

Vladimir Fock

Fock space is of great value when particles are created and annihilated. For example, in a completely different energy regime, a highly energetic electron that scatters at many MeV might exit the scattering region having produced an electron-positron pair. In this case, the number of massive particles is not fixed. Fock space is also extremely useful in other scientific areas, where it serves as a bookkeeping tool. Cases that come to mind are condensed matter and electronic structure.[8]

The above considerations give cause to ponder regarding what might be anticipated in extensions of the results obtained so far. Phonons emerged when Q_k and P_k were quantized using commutator relations. The lattice model that was enlisted at the outset, though unconventional, is transparent. Every step along the way, from eqn (1.1) onward, could be followed. Of course for real crystals N is huge. Consequently, adjacent k values are so close to one another in energy that individual k values are unobservable except in unusual circumstances. The only one I know of is crystalline hydrogen at ultra-low temperatures. In other words, individual k values are difficult to resolve experimentally.

To see how close in energy adjacent levels might be to one another, consider a 5 mm crystal length and a lattice spacing of 0.5 nm. In this case, 10^7 sites are encountered along a straight line that extends over the 5 mm length. The energy separation for adjacent k values for an assumed Ω of 10 cm^{-1} ($\Omega_0 = 0$) is $\sim 10^{-5}$ cm^{-1}. For such small spacing, the lattice modes comprise a dense quasi-continuum that might as well, for spectroscopic purposes, be a true continuum.

---

[8] A Fock space is comprised of Hilbert spaces in a way that facilitates the introduction and removal of particles (field quanta), including their statistical property (fermion or boson). It is a direct sum of Hilbert spaces, each of which is a tensor product of single particle Hilbert spaces for a given number of particles. The direct sum is taken over numbers of particles. This makes it useful in areas like electronic structure theory, where using Fock space to put electrons into orbitals results in properly anti-symmetrized (with respect to particle exchange) wave functions.

The number of single particle Hilbert spaces is not limited except trivially: for fermions there cannot be more particles than there are spin-orbitals in each single particle Hilbert space. In this case, the dimension of the Fock space is 2^M, where M is the number of orbitals. For example, suppose there are four orbitals (numbered 1, 2, 3, 4) and we wish to fill the Fock space, *i.e.*, include all possibilities. The case of no particles in any of the orbitals is not exciting, but it cannot be overlooked. This counts as 1 arrangement. One particle can be put into any of the orbitals (4 arrangements). Two particles can be introduced as follows: 12, 13, 14, 23, 24, 34 (6 arrangements). Three particles can be introduced as follows: 123, 124, 134, 234 (4 arrangements). Four particles have 1 arrangement. In total there are $1 + 4 + 6 + 4 + 1 = 16 = 2^4 = 2^M$ arrangements.

Despite this seemingly continuous nature of the k distribution, the shape of the dispersion curve, such as the one shown in Fig. 8, is preserved, with the edge of the first Brillouin zone occurring when a half wavelength is equal to the lattice spacing a, namely, $k = \pi / a$. In other words, the edge of the Brillouin zone, and related phenomena, are manifestations of the spacing between sites. This is independent of the length of the crystal, which determines the energy spacing for adjacent k values.

Now imagine that the spacing a between adjacent lattice sites is made progressively smaller. The mass density, $\rho = m / a$, is forced to remain constant throughout this progression, so as a gets smaller, m gets smaller in concert with a. We are simply replacing the discrete masses m with a uniform mass density. In the limit where m and a each approach zero, but with a constant value of $\rho = m / a$, the system is described using a mass density, or *mass field*. Waves still propagate, so the phonon description has not been compromised, but there is no longer a first Brillouin zone, at least a pragmatic one. The edge at $k = \pi / a$ now lies at $k = \infty$.

It is noteworthy that the quantization of a massive field using commutators yields bosons as the field quanta. As we shall see later, the massless electromagnetic field is also quantized using commutators, in this case yielding bosons that are called photons. In electromagnetism we deal with a *vector field*. A consequence of the vector nature is that the photon has an intrinsic spin of one. On the other hand, a phonon, being the quantum of a scalar field, has a spin of zero.

These rules about intrinsic spins can be obtained using an important theorem that was introduced about a century ago by the brilliant German mathematician Emmy Noether. Of course, it is referred to as Noether's theorem. The photo was taken slightly before she was driven from Germany by the Nazis because she was Jewish. Thus, phonons and photons are each bosons, even though their fields are of quite different character: massive scalar versus massless vector, respectively. It turns out that quantization of a *spinor field* can only be carried out using anti-commutators. As a result, the quanta of the spinor field are fermions. They obey the Pauli exclusion principle.

Emmy Noether

The next three chapters are devoted to: (i) the ionic lattice, (ii) thermodynamic properties such as heat capacity and thermal conductivity, and (iii) infrared spectroscopy. This connects well with the electronic structure of periodic lattices that we encountered in Part A: band structure, Brillouin zones, metals and insulators, and so on. Following this, we shall return to our toy model and extend it. This is carried out in steps. First, the aforementioned transition from the discrete lattice to a continuous mass field is carried out. The resulting displacement and momentum fields, $\hat{q}(x)$ and $\hat{p}(x)$, are then expanded in a basis of their k-space Fourier complements, $\hat{Q}_k$ and $\hat{P}_k$, as was done with the discrete lattice. This enables the Hamiltonian to be expressed as the spatial integration of a *Hamiltonian density*.

Extension to a continuous mass distribution yields a straightforward expression for the displacement field, which varies continuously with x. It is of utmost significance that the k-space remains discrete, which simplifies the math considerably. In the sections that fol-

low: the time dependence is determined, results are extended to 3D, and then to all of spacetime. In the end we arrive at a profound result – all the more impressive considering the humble model from which it originated. It is possible with little additional work to extend this result to the relativistic regime and obtain the Klein-Gordon equation for spin-zero massive scalar fields.

7. The Ionic Lattice

A popular model is that of an ionic lattice, for example, an alkali halide such as NaCl or an alkaline earth oxide such as MgO. We shall work through this in 1D, as it is easy and serves as a basis for a discussion of properties such as thermal conductivity and the infrared absorption spectra of such systems.

Figure 11 depicts the 1D ionic lattice. It is assumed that the masses indicated with red circles, m_A, are heavier than those indicated with blue circles, m_B. The yellow boxes indicate 1D unit cells. Because the lattice is ionic, each nearest-neighbor equilibrium spacing has the same value, a. This aspect – a single value of nearest neighbor equilibrium spacing– would not appear in a network in which there is covalent character. Because the repeat unit is $2a$, the edges of the first Brillouin zone are at $\pm \pi/2a$.

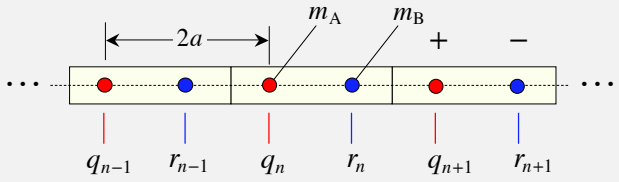


Figure 11. 1D ionic lattice: The repeat distance is $2a$ so the edges of the first Brillouin zone are at $\pm \pi/2a$. The lattice could be NaCl, MgO, or any other ionic lattice. Red masses (m_A) are heavier than blue masses (m_B).

As in Section 1, it is straightforward to write the equations of motion and the Hamiltonian. We need not concern ourselves with second quantization, as we have seen how it is implemented, and this does not change. Thus, our ionic lattice model will be solved using the fastest route: (i) equations-of-motion, (ii) characteristic equation, and (iii) the answer. The term "solved" means that the dispersion relation (ω_k versus k) is obtained. It can then be interpreted using the material in Section 6: *Interpretation*.

To obtain the dispersion relation and normal mode displacements, assume sinusoidal time variation. It does not matter whether $\sin\omega_k t$, $\cos\omega_k t$, $e^{i\omega_k t}$, or $e^{-i\omega_k t}$ is used, as each gives the same result. Namely, differentiating displacement q_n twice with respect to time gives $\ddot{q}_n = -\omega_k{}^2 q_n$. Likewise, for the displacement r_n we have $\ddot{r}_n = -\omega_k{}^2 r_n$. Thus, the equations-of-motion yield

$$m_A \omega_k{}^2 q_n = \kappa\left(2q_n - r_{n-1} - r_n\right) \tag{1.48}$$

$$m_B \omega_k{}^2 r_n = \kappa\left(2r_n - q_n - q_{n+1}\right), \tag{1.49}$$

where κ is the nearest neighbor harmonic oscillator force constant, and m_A and m_B are the masses associated with displacements q and r, respectively (Fig. 11). The idea is to obtain relative displacement amplitudes, including phase. Therefore, eqns (1.48) and (1.49) are expressed in terms of the complex amplitudes A and B: $q_n = A e^{i2kan}$ and $r_n = B e^{i2kan} e^{ika}$, where q_n and r_n are real. In r_n, the factor e^{ika} that is appended is not essential, but proves mathematically expedient. Substituting these into eqns (1.48) and (1.49), and carrying out some algebra (box below), yields

$$\begin{pmatrix} \omega_k^2 - \dfrac{2\kappa}{m_A} & \dfrac{2\kappa \cos ka}{m_A} \\ \dfrac{2\kappa \cos ka}{m_B} & \omega_k^2 - \dfrac{2\kappa}{m_B} \end{pmatrix} \begin{pmatrix} A \\ \\ B \end{pmatrix} = 0. \tag{1.50}$$

Substitution of $q_n = A e^{i2kan}$ and $r_n = B e^{i2kan} e^{ika}$ into eqns (1.48) and (1.49) yields

$$m_A \omega_k^2 A e^{i2kan} = \kappa \left(2A e^{i2kan} - B e^{i2ka(n-1)} e^{ika} - B e^{i2kan} e^{ika} \right)$$

$$m_B \omega_k^2 B e^{i2kan} e^{ika} = \kappa \left(2B e^{i2kan} e^{ika} - A e^{i2kan} - A e^{i2ka(n+1)} \right)$$

Dividing the first equation by e^{i2kan} / m_A and the second equation by $e^{i2kan} e^{ika} / m_B$ yields

$$\omega_k^2 A = \frac{2\kappa}{m_A} A - \frac{\kappa (e^{-ika} + e^{ika})}{m_A} B = \frac{2\kappa}{m_A} A - \frac{2\kappa \cos ka}{m_A} B$$

$$\omega_k^2 B = \frac{2\kappa}{m_B} B - \frac{\kappa (e^{-ika} + e^{ika})}{m_B} A = \frac{2\kappa}{m_B} B - \frac{2\kappa \cos ka}{m_B} A$$

In matrix form, this is eqn (1.50).

To obtain the eigenvalues, ω_k^2, the determinant of the matrix in eqn (1.50) is set equal to zero. Some algebra yields

$$\omega_k^2 = \kappa \left(\frac{1}{m_A} + \frac{1}{m_B} \right) \pm \kappa \left(\left(\frac{1}{m_A} + \frac{1}{m_B} \right)^2 - \frac{4}{m_A m_B} \sin^2 ka \right)^{1/2}. \tag{1.51}$$

This is the dispersion relation. Only the positive square root of $(\omega_k^2)^2$ is used. On the right hand side of eqn (1.51) the plus and minus signs in front of the radical correspond to the optical and acoustic branches, respectively. A sketch (not to scale) is given in Fig. 12. This result will now be interpreted.

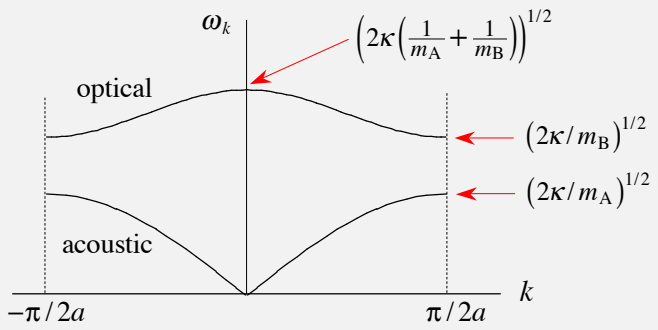


Figure 12. Dispersion relation for a 1D ionic lattice with different masses ($m_B < m_A$): note that the zone edges lie at $\pm\pi/2a$, rather than $\pm\pi/a$.

Amusement

An instructive exercise is to set the masses equal: $m_A = m_B = m$. The fact that the lattice is ionic is reflected in the fact that a single spring constant κ enters. Otherwise, it has no bearing on the dispersion relation given by eqn (1.51). Except for the size of the unit cell ($2a$) and the coupling strengths, the lattice might just as well be comprised of Ar atoms as of Na^+ and Cl^-, or of Mg^{+2} and O^{-2}.

The infrared spectrum of an ionic lattice is another matter, one we will get to soon. Right now, we anticipate that when the positive and negative ions each have the same mass the result for a monatomic lattice will follow, despite the fact that the ions are assumed to be distinct. Think of NaF, KCl, RbBr, and CsI. The atomic masses are not so much different with these pairs: Na = 23 and F = 19; K = 39 and Cl = 35.5; Rb = 85.5 and Br = 80; Cs = 133 and I = 127.

It is interesting to see how the case $m_A = m_B = m$ develops mathematically, starting from eqn (1.51) and Fig. 12. Let us work through this one step at a time. When $m_A = m_B = m$ is introduced into eqn (1.51), a simple expression for ω_k^2 is obtained:

$$\omega_k^2 = \frac{2\kappa}{m}\left(1 \pm \cos ka\right). \tag{1.52}$$

Let us now see what this looks like and how it is to be interpreted. The acoustic branch [minus sign in eqn (1.52)] is $\omega_k^2 = (2\kappa/m)(1-\cos ka) = (4\kappa/m)\sin^2(ka/2)$. Thus, ω_k has values of zero at $k = 0$, and $(2\kappa/m)^{1/2}$ at the zone edges ($\pm\pi/2a$). Something is clearly amiss, as the group velocity does not vanish at the zone edges.

Recall that the dispersion relation for the acoustic branch of the model introduced at the beginning of Section 1 is given by: $\omega_k^2 = (4\kappa/m)\sin^2(ka/2)$, *i.e.*, eqn (1.39) with $\Omega^2 = \kappa/m$ and $\Omega_0 = 0$. In other words, eqn (1.39) for ω_k^2 looks the same as the expression for ω_k^2 in the above paragraph. However, at the Brillouin zone edge of the Section 1 model ω_k is equal to $(4\kappa/m)^{1/2}$, which is $2^{1/2}$ larger than the result obtained using eqn (1.52). The difference is due to the zone edges: $\pm\pi/a$ with the Section 1 toy

model, and $\pm\pi/2a$ with eqn (1.52), which follows directly from eqn (1.51). Now that we have identified where the disparity originates, it remains to reconcile the difference.

The choice of + in eqn (1.52) gives the dispersion relation for the optical branch: $\omega_k{}^2 = (2\kappa/m)(1+\cos ka) = (4\kappa/m)\cos^2(ka/2)$. Thus $\omega_k = 2(\kappa/m)^{1/2}$ when $k = 0$, and at the zone edge it is $(2\kappa/m)^{1/2}$. We see that the acoustic and optical branches meet at the zone edges defined by eqn (1.52), as shown in Fig. 13. It was mentioned in the paragraph below eqn (1.52) that the group velocity is not zero at the zone edges, which means that something is wrong. In fact, we now realize that the whole idea of optical and acoustic branches in this system makes no sense. There is no optical branch in a monatomic lattice!

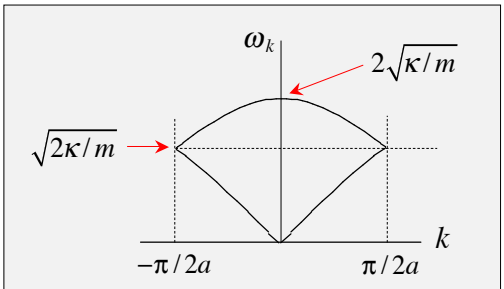
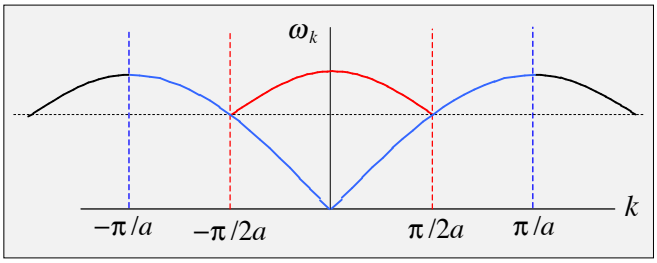


Figure 13. Using equal masses with eqn (1.51) yields this dispersion relation.

When care is taken with the zone edges the riddle is solved. Figure 14 illustrates what is going on. The so-called optical branch in eqn (1.52) and Fig. 13 is shown in red in Fig. 14. We see that it is a displaced part of the acoustic branch. When it is moved to its right and left by π/a, the blue curve is obtained. The blue curve between the blue vertical dashed lines is the correct acoustic branch of the monatomic lattice.

Figure 14. The region between dashed red lines is from Fig. 13. The red curve can be displaced horizontally by π/a without affecting anything physical. This shows that the optical branch evolves into part of the acoustic branch as masses become equal.



Displacements

The above example of a single value for the ion masses underscores the fact that care needs to be taken when interpreting dispersion relations. Let us now return to the case of different masses for the positive and negative ions. The goal is to see how ion displacement varies with k throughout the first Brillouin zone for the acoustic and optical branches. The ratio B/A is obtained by using the characteristic equation whose matrix form is given by eqn (1.50). This yields

$$\frac{\mathrm{B}}{\mathrm{A}} = \frac{2\kappa - m_{\mathrm{A}}\omega_k^2}{2\kappa\cos ka} = \frac{2\kappa\cos ka}{2\kappa - m_{\mathrm{B}}\omega_k^2}. \tag{1.53}$$

These expressions enable us to examine displacements for the acoustic and optical branches as a function of location within the Brillouin zone.

Acoustic Branch

The dispersion relation of the acoustic branch in the $k \sim 0$ (long wavelength) region is linear, as seen in Fig. 12. If a small ka value is used with eqn (1.51), the proportionality between ω_k and k is obtained. From eqn (1.51), we see that the small-ka limit yields

$$\omega_k^2 = \kappa\left(\frac{1}{m_A}+\frac{1}{m_B}\right) \pm \kappa\left(\left(\frac{1}{m_A}+\frac{1}{m_B}\right)^2 - \frac{4}{m_A m_B}k^2a^2\right)^{1/2}. \tag{1.54}$$

Choosing the minus sign and introducing the reduced mass: $m_{AB} = m_A m_B / m_T$, where m_T is the total mass of the ion pair ($m_T = m_A + m_B$), yields

$$\omega_k^2 = \frac{\kappa}{m_{AB}} - \kappa\left(\left(\frac{1}{m_{AB}}\right)^2 - \frac{4k^2a^2}{m_A m_B}\right)^{1/2} \tag{1.55}$$

$$= \frac{\kappa}{m_{AB}}\left(1-\sqrt{1-\frac{4(m_{AB})^2}{m_A m_B}k^2a^2}\right) \tag{1.56}$$

$$\approx \left(\frac{2\kappa}{m_T}a^2\right)k^2. \tag{1.57}$$

Thus, we have the speed associated with the linear part of the acoustic branch:

$$\boxed{\omega_k = v_0 k \quad \text{where } v_0^2 = 2\kappa a^2 / m_T.} \tag{1.58}$$

In the above equations v_0 is the speed of sound in the medium. Note that when the ion masses are equal ($m_A = m_B = m$) we have $v_0 = (\kappa/m)^{1/2}a$.

In the low-k regime, acoustic wave packets do not distort as they propagate with speed v_0. Typical v_0 values for ionic crystals are $\sim 5000\ m/s$ (for example, 4730 for NaCl and 4950 for LiF). To put this on a molecular time scale, at this speed an acoustic disturbance travels a distance of 5 nm in 1 ps.

Referring to eqn (1.53), for the acoustic branch in the low-k regime, the ratio B/A is close to unity. In other words, the displacements of the positive and negative ions have approximately the same magnitude. Moreover, because of the long wavelength, the ions oscillate in phase, as indicated in Fig. 15(a). If one peruses the entire length of the linear lattice, the spatial variation of the displacements can be seen in the low-k regime. However, they cannot be seen on the much smaller scale used in Fig. 15 of just a few sites.

As k increases in magnitude in the acoustic branch, B/A decreases. In other words, the lighter ions experience progressively smaller excursions from equilibrium than do the heavier ions. At the zone edges, displacements of the lighter ions are zero. The heavier ions move while the lighter ions remain stationary. The phase factor e^{i2kan} is equal to $(-1)^n$, so displacement changes direction from one heavy ion to the next, as shown in Fig. 15(b).

You can tell by just looking at Fig. 15(b) that the group velocity is zero at the edge of the Brillouin zone. There is simply no net transport of energy because the lattice vibration is a standing wave. This can also be calculated using eqn (1.51).

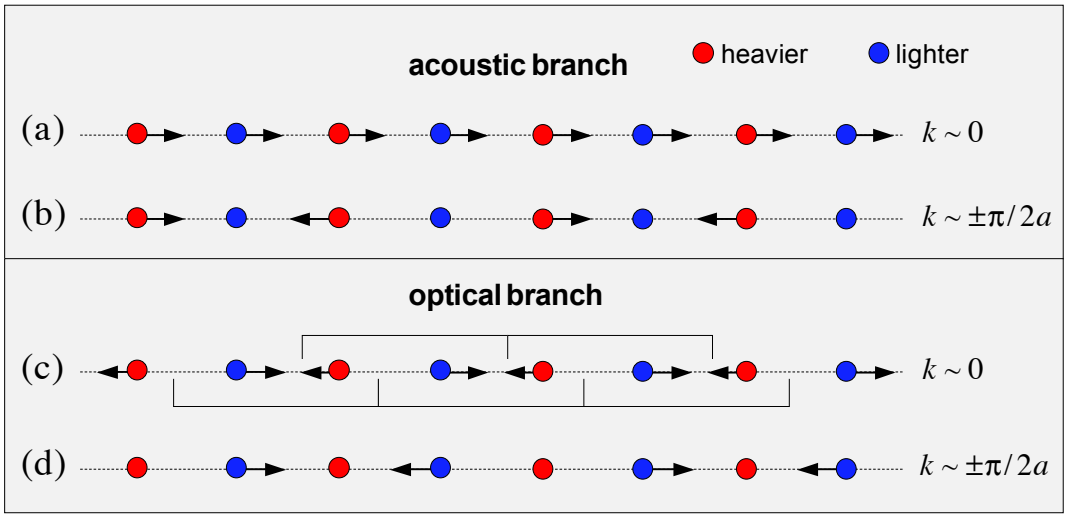


Figure 15. Displacements for acoustic and optical modes in the regions $k \sim 0$ and $k \sim \pm\pi/2a$: red and blue represent the heavier and lighter ions, respectively. (a) The $k \sim 0$ part of the acoustic branch has almost equal displacement for each ion, regardless of whether it is the positive or negative ion. (b) At the zone edges ($k \sim \pm\pi/2a$), a standing wave consists of the heavier ions moving while the lighter ions remain still. (c) The region near $k = 0$ consists of vibrations about the centers-of-mass of ion pairs. Locations of these centers-of-mass are indicated using one or the other of the combs. Energy ceases to propagate as $k = 0$ is approached; vibrations are local. (d) At the zone edge the heavier ions remain still while the lighter ions oscillate. Again, energy does not propagate. Referring to Fig. 12, notice that the group velocity $v_g = \partial\omega/\partial k$ vanishes at $k = 0$ and $\pm\pi/2a$ for the optical branch.

Optical Branch

The optical branch can be understood in a like manner by using Fig. 12 and eqns (1.51) and (1.53). In the $k \sim 0$ long wavelength regime, we see from Fig. 12 that the phase velocity, $v_p = \omega/k$, approaches infinity as k approaches zero, whereas v_g approaches zero as k approaches zero. This is a quite remarkable difference: one velocity goes to infinity while the other goes to zero! You might guess that for $v_g \to 0$ an analogous situation will arise as happened with the acoustic branch at $k = \pm\pi/2a$, where a standing wave was identified. Let us see how this works.

Plugging the $k = 0$ value of $\omega_k{}^2$ for the optical branch into eqn (1.53) yields a B/A ratio of $-m_A / m_B$. Figure 15(c) shows what this looks like. Each pair of positive and negative ions oscillates about its center-of-mass. The centers-of-mass of ion pairs are indicated with the combs. Two combs are shown because defining ion pairs is arbitrary. That is, going from left to right, the first and second ions (*i.e.*, red followed by blue) can be paired, or, alternatively, the second and third (blue followed by red) can be paired. Again we have a standing wave, so $v_g = 0$.

An intuitive parallel can be drawn with CO_2 normal modes: ν_1 symmetric stretch and ν_3 asymmetric stretch. With ν_1, the two oxygen atoms oscillate out of phase with respect to one another, *i.e.*, in opposite directions with the carbon atom stationary. With ν_3, the two oxygen atoms move in the same direction, and the carbon atom moves in the direction opposite that of the oxygen atoms. All three atoms moving in the same direction is overall translation.

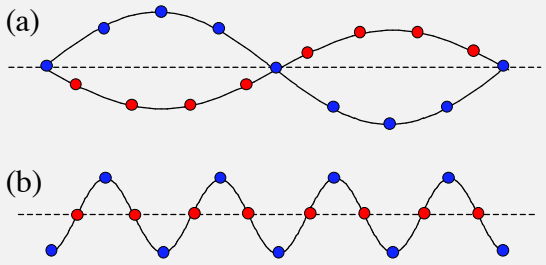


Figure 16. TO phonons illustrate optical branch displacements. (a) For intermediate k values the lighter ions (blue) have larger displacements than the heavier ions (red). (b) At the zone edge the heavier ions remain stationary while the lighter ions oscillate up and down.

Referring to Fig. 12, at the band edge we have $v_g = 0$, so a standing wave is anticipated. In this case the positive radical is used with eqn (1.51). The expression for the radical simplifies because $\sin^2 ka = 1$, in which case the radical becomes: $+\kappa(m_B{}^{-1} - m_A{}^{-1})$. Consequently, $\omega_k = (2\kappa / m_B)^{1/2}$, as indicated in Fig. 12. When this value of ω_k is put into eqn (1.53), we find that the ratio B/A is zero. In other words the heavier ion is stationary while the lighter ion carries all of the motion.

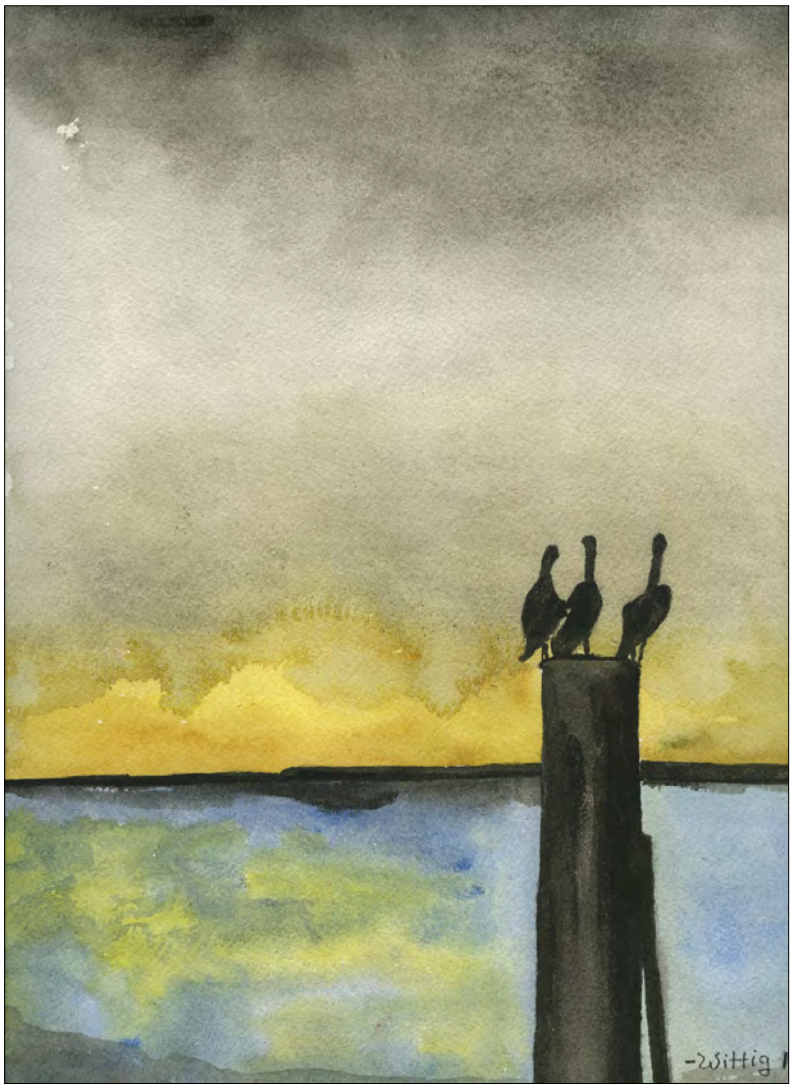
In the region where $|k|$ is intermediate between 0 and $\pi/2a$, it is easier to visualize the displacements by using the transverse optical branch of a 3D model than the longitudinal optical branch of the 1D model. Figure 16(a) shows displacements for a TO phonon. The zone edge version is shown in Fig. 16(b).

Chapter 2.

Heat Capacity and Thermal Conductivity

Wash Day
Robert Wittig

Contents



Birds of Lill
Robert Wittig

It is inexcusable for scientists to torture animals; let them make their experiments on journalists and politicians.

Henrik Ibsen

2.1. Preliminary Comments

To determine a system's thermodynamic properties, it is necessary to know how its stored (non-chemical) energy is distributed over the degrees of freedom of its particles and quasiparticles at a given temperature. In this chapter, we shall be concerned with the heat capacity and thermal conductivity of non-magnetic, electrically insulating crystals. In these systems, phonons are solely responsible for heat capacity and thermal conductivity, with the exception of very low temperatures (say ~ 1 K), where impurities, defects, and boundaries limit thermal conductivity.

Metals are different. Their ionic lattices give rise to phonons, including low frequency ones that result in good heat capacity. In fact, phonons play the dominant role in the heat capacity and thermal conductivity of metals at all but the lowest temperatures. Electrons play a significant role only at quite low temperatures. The reason electrons contribute so little is due to the fact that they are fermions. Therefore, only electrons that have energy near the Fermi level can move about freely and participate in heat capacity. The rest of the electrons are, so to speak, frozen out. The reason they are important at low temperatures is that the phonon contributions freeze out more rapidly than the electron contributions as temperature is lowered.

There is nothing so complicated that it cannot, with sufficient effort, be made more complicated.

To keep matters simple, the electronic contribution to the heat capacity and the thermal conductivity will not be considered. In other words, we shall deal with non-magnetic, electrical insulating crystals or, it can be said, the phonon contribution to the heat capacity of a metal.

In this chapter, we shall begin by using the distribution function for a canonical ensemble of bosons to obtain the average number of quanta in a given mode. Following this, elementary models of heat capacity and thermal conductivity will be introduced. More detailed models would require more time than we have available for these topics. It is hoped that the introductory treatment presented below will open the door to related topics, should you wish to pursue the subject further.

2.2. Canonical Ensemble of Bosons

The average number of quanta in a given mode is calculated here for a sample in thermal equilibrium at temperature T. This is done first for a 1D system, and the result is then extended to 3D. The 2D case is assigned as an exercise. The Boltzmann factor e^{-E/k_BT} is used to obtain the average energy of the k^{th} mode, which shall be denoted $\langle E_k \rangle$. It is not necessary to include the zero point energy unless chemical bonds are broken or phase transitions take place. These phenomena will not be considered here, so zero point energy

shall be suppressed. Also, in discussing heat capacity, it is assumed for the time being that all vibrations are harmonic. That is, anharmonicity is suppressed. However, you will see that anharmonicity proves to be crucial when it comes to thermal conductivity, and we will discuss its role there. It is also responsible for thermal expansion, phonon removal rates, and so on.

Because the energies available to the k^{th} mode are integer multiples, n_k, of the energy quantum $\hbar\omega_k$, $\langle E_k \rangle$ is related trivially to the average number of quanta, $\langle n_k \rangle$:

$$\langle E_k \rangle = \hbar\omega_k \langle n_k \rangle = Z^{-1} \sum_{n=0}^{\infty} n\hbar\omega_k \exp\left(-\frac{n\hbar\omega_k}{k_B T}\right). \tag{2.1}$$

Explicit $\hbar$ will be used for a while; the Boltzmann constant is k_B; and the partition function, *Z*, is given by

$$Z = \sum_{n=0}^{\infty} \exp\left(-\frac{n\hbar\omega_k}{k_B T}\right). \tag{2.2}$$

Notational simplification is achieved by using $x = \exp(-\hbar\omega_k / k_B T)$, in which case eqn (2.1) gives

$$\langle n_k \rangle = Z^{-1} \sum_{n=0}^{\infty} n x^n = Z^{-1}\left(x + 2x^2 + 3x^3 ...\right). \tag{2.3}$$

Thus, eqn (2.2) becomes

$$Z = \sum_{n=0}^{\infty} x^n = 1 + x + x^2 ... = 1 + x(1 + x + x^2 ...) = 1 + xZ, \tag{2.4}$$

yielding a compact expression for the partition function:

$$Z = \frac{1}{1-x}. \tag{2.5}$$

The parenthetic term in eqn (2.3) can be written $x(1 + x + x^2 ...)^2$, which is identified as being equal to xZ^2. Thus, eqn (2.3) becomes

$$\langle n_k \rangle = xZ = \frac{x}{1-x}. \tag{2.6}$$

Finally, replacing *x* with $\exp(-\hbar\omega_k / k_B T)$ yields the standard form:

$$\langle n_k \rangle = \frac{1}{e^{\hbar\omega_k / k_B T} - 1}. \tag{2.7}$$

This is the average number of quanta excluding zero point for a thermal distribution of phonons whose energy quantum is $\hbar\omega_k$. It is the Bose-Einstein distribution for bosons at temperature *T*. The fact that an average value appears in eqn (2.7) does not mean that $\langle n_k \rangle$ is large. If $\langle n_k \rangle$ is small, the assumption of a canonical distribution gives a correct average value, but it turns out that n_k fluctuates from one *k* value to the next. However, the mode density is usually large, in which case a smooth distribution of the number of modes versus energy can be recovered by averaging over some number of modes according to an *n*-point smoothing routine.

Referring to eqn (2.7), the corresponding expression for fermions differs in a way that is subtle and of enormous significance. The -1 is replaced by $+1$, and $\hbar\omega_k$ is replaced by $E - E_F$, where *E* is the energy of a single-particle state and E_F is the Fermi energy. It is the energy of the highest occupied state at 0 K.[1]

$$f(E) = \frac{1}{e^{(E-E_F)/k_B T} + 1}. \tag{2.8}$$

The difference between eqns (2.7) and (2.8) is profound in the extreme. For example, the high temperature limit of eqn (2.7) is $k_B T/\hbar\omega_k$, whereas the high temperature limit of eqn (2.8) ($k_B T >> E - E_F$) is ½. The low temperature limit of eqn (2.7) is $e^{-\hbar\omega_k / k_B T}$, whereas the low temperature limit of eqn (2.8) is 1 for $E - E_F < 0$, and 0 for $E - E_F > 0$. It is not possible for two or more identical fermions (electrons) to occupy a given single-particle state (spin-orbital), whereas phonons can occupy a given mode with as many quanta as possible, limited only by degeneracy and the Boltzmann factor. Figure 1 illustrates this difference.

We saw in Chapter 1 that the quantization of a field of harmonically coupled mass points yields phonons. These are massless bosons. There is no energy involved in their creation and annihilation other than the energy of their quanta. On the other hand, fermions have mass, so at the low energies of concern here we do not deal with their creation and annihilation. These processes take place at high energy because the field quanta (the fermions themselves) have mass whose energy is mc^2. In other words, the number of fermions is assumed to remain constant, obviating the need for a field theoretic description of their dynamics.

---

[1] Equation (2.8) is not correct, though it is an extremely accurate approximation for the systems and temperatures of concern. The term E_F that appears in eqn (2.8) should be the chemical potential, μ, which is also referred to as the Fermi *level*. Only at much higher temperatures than those of interest here does μ differ significantly from E_F. The Fermi energy is easier to visualize than the Fermi level. The latter is discussed in Chapter 3.

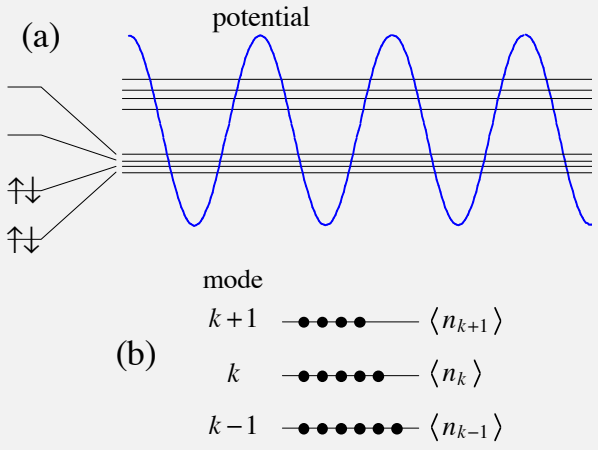


Figure 1. (a) Electrons in a 1D periodic potential occupy orbitals according to the Pauli exclusion principle: one per spin orbital. For a given temperature the distribution function is given by eqn (2.8). (b) A given phonon mode (k value) has an energy quantum $\hbar\omega_k$. Many phonons (dots) can occupy the same mode. The average number of phonons per mode is given by eqn (2.7)

The bottom line is that fermions obey the Pauli exclusion principle, which sets their thermodynamic properties apart from those of bosons and other massless quasiparticles.

2.3. Heat Capacity

Despite the fact that the heat capacity of an electrically insulating crystal is conceptually straightforward, a first-principles calculation for a real crystal would be too arduous to be undertaken here. To strike a compromise, the underlying principles will be presented, and then a simple model due to Debye will be introduced and discussed. This model accounts for the main features of heat capacity, and it is applicable from the lowest to the highest temperatures. It is generally in reasonable agreement with experiment. This is not surprising given that the Debye model is an interpolation between the low and high temperature limits, each of which is on sound theoretical footing. These limiting cases are easy to derive and understand, and they will be discussed.

A good starting point is eqn (2.7). The heat capacity of a substance is the rate of change of its stored energy with temperature. In other words, if a 1 K temperature increase results in the stored energy of one mole of the substance increasing by Q Joules, the heat capacity C is equal to $\mathrm{Q\ Joules \cdot K^{-1} \cdot mole^{-1}}$. For a crystal, there is little difference between the heat capacities at constant pressure and at constant volume, C_P and C_V, respectively, because little work is done by thermal expansion of the crystal. Thus, we will drop the P and V subscripts and refer to the heat capacity simply as C. For a single mode (single k value), the single-mode heat capacity C_k is obtained straightaway by using eqns (2.1) and (2.7). The result is

$$C_k = \frac{\partial \langle E_k \rangle}{\partial T} = \hbar\omega_k \frac{\partial}{\partial T}\left(\frac{1}{e^{\Theta_k/T} - 1}\right), \tag{2.9}$$

where $\Theta_k = \hbar\omega_k / k_B$. Carrying out the differentiation yields

$$C_k = k_B \left(\frac{\Theta_k}{T} \right)^2 \frac{e^{\Theta_k / T}}{(e^{\Theta_k / T} - 1)^2} . \tag{2.10}$$

Notice that this result, which was easy to obtain, has nothing to do with the dimension of the space. It can be applied to 1D, 2D, and 3D models. The dimension of the space enters when we include the participating frequencies, which is done according to the dispersion relation. This is easy with 1D models, but it can prove challenging with 3D models. Again, we shall start with 1D and proceed to 3D.[2]

The *k*-space density, $\rho(k)$, is the number of modes per unit wave vector. Sometimes this is referred to as the density of states in *k*-space. This usage will be avoided, because we shall take the term "state" in the context of density of states to mean a fully resolved (non-degenerate) state that has a specified energy.

The *k*-space density $\rho(k)$ is now calculated and it is used to obtain the density of mode frequencies, $g(\omega)$. This is where the dispersion relation enters. The average energy is obtained by combining $g(\omega)$ with the average energy per mode. We shall start with $\rho_1(k)$ (the subscript denotes 1D) and the dispersion relation for one of the 1D models introduced earlier. The dispersion relation enables us to convert $\rho_1(k)$ to $g_1(\omega)$, the number of corresponding modes per unit frequency interval. In other words, $\rho_1(k)$ is mapped onto $g_1(\omega)$ using the relationship between *k* and ω given by the dispersion relation. Hereafter, the convention $\hbar = 1$ will be used unless otherwise noted.

2.4. One-Dimensional Monatomic Lattice

The *k*-space density for a 1D monatomic lattice, $\rho_1(k)$, is equal to $L/2\pi$, as indicated in Fig. 2. The dispersion relation that will be used in the material that follows is $\omega^2 = 4\Omega^2 \sin^2(ka/2)$. For positive frequencies, this dispersion relation is $\omega = 2\Omega \sin(ka/2)$. Henceforth, the *k* subscript that appeared earlier on ω_k and C_k is dropped, replaced by $\omega(k)$ and $C(k)$, as all distributions in *k* and ω spaces will be treated as continuous.

[2] It is not problematic in the present context, but it is often the case that one should be cautious when referring to the dimension of a space. It would usually be more appropriate to replace the word dimension with direction. The reason for the distinction is that the spin of an elementary fermion such as the electron requires a 3D space. In 2D there is no spin-statistics connection of the kind that gives the Pauli exclusion principle. Frank Wilczek worked this out and assigned the term anyons to such elementary 2D particles.

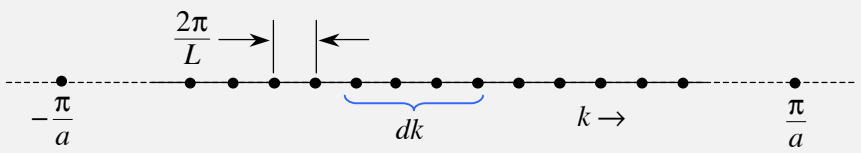


Figure 2. Allowed k values are indicated with black dots. For a 1D monatomic lattice, the separation between adjacent k-values is equal to $2\pi / L$. In a small interval dk, the number of k values is equal to $\rho_1(k)dk$, where $\rho_1(k)$ is the density of k values. It is necessary to satisfy $dk >> 2\pi / L$ if we are to have an effectively continuous k-space density $\rho_1(k)$. Note that integration of $\rho_1(k)dk$ over the range $-\pi / 2 \le k \le \pi / 2$ yields the number of modes, N.

To proceed, the fact that a given number of modes can be expressed in either ω-space or k-space, that is, $g_1(\omega)d\omega = \rho_1(k)dk$ (see Fig. 3), is used to write

$$g_1(\omega) = \rho_1(k)\left(\frac{dk}{d\omega}\right). \tag{2.11}$$

We shall first obtain $dk / d\omega$ and then combine this with $\rho_1(k)$ to obtain $g_1(\omega)$.

For spaces of any dimension (1D, 2D, or 3D), when obtaining $g(\omega)$ it is understood that the k that appears in the k-space density $\rho(k)$ is to be replaced with a function of ω. Specifically, k is replaced with $k(\omega)$ according to the dispersion relation. In the 1D case under consideration, $\rho_1(k)$ is constant ($L/2\pi$), so this substitution does not arise. We simply replace $\rho_1(k)$ with $L/2\pi$. The next section deals with a 3D case, where the issue is revisited because $\rho_3(k)$ is proportional to k^2. Note that the parenthetic term on the right hand side of eqn (2.11) is the inverse of the group velocity: $v_g \equiv \partial\omega / \partial k$.

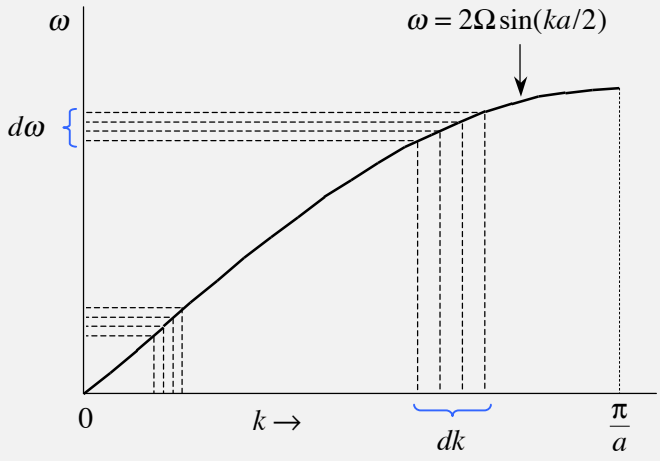


Figure 3. The interval dk is mapped onto the interval $d\omega$ through the dispersion relation given at the top. As the edge of the Brillouin zone is approached along the dispersion curve, the slope flattens and therefore the density of ω values increases. At low k, the linear relationship between ω and k (dashed lines in lower left) gives the speed of sound, $\omega = vk$.

Differentiation of the dispersion relation $\omega = 2\Omega\sin(ka/2)$ yields

$$\frac{dk}{d\omega} = \frac{1}{\Omega a\cos(ka/2)}. \tag{2.12}$$

Putting this into eqn (2.11), and using $\rho_1(k) = L/2\pi$, yields

$$g_1(\omega) = \frac{L}{2\pi}\frac{1}{\Omega a\cos(ka/2)}. \tag{2.13}$$

As usual, k spans the $2\pi/a$ range that is symmetric about $k = 0$: $-\pi/a \le k \le \pi/a$. Keep in mind that the k that appears on the right hand side of eqn (2.13) is to be replaced with a function of ω according to the dispersion relation. Writing the right hand side of eqn (2.13) in terms of ω is generally awkward and offers no clear advantage. Thus, we will stick with eqn (2.13) as it stands, with the understanding that the ω dependence of the right hand side is obtained through the dispersion relation $\omega = 2\Omega\sin(ka/2)$.[3] To obtain $g_1(\omega)$ in terms of only positive k values, simply multiply the right hand side of eqn (2.13) by two.

The expression for $g_1(\omega)$ given by eqn (2.13) has the interesting feature that it goes to infinity when $\cos(ka/2) = 0$ at the edge of the first Brillouin zone. However, there is no cause for alarm. We know that this mathematically singular behavior cannot result in something strange happening, because in 1D the number of modes is equal to N. In other words, integrating $g_1(\omega)d\omega$ and $\rho_1(k)dk$ must yield N in both cases. The singularity that appears in eqn (2.13) is said to be integrable.[4]

---

[3] In the present example, an expression for $\cos(ka/2)$ is readily obtained from the dispersion relation $\omega = 2\Omega\sin(ka/2)$:

$$\left(\frac{\omega}{2\Omega}\right)^2 = \sin^2(ka/2) = 1-\cos^2(ka/2) \implies \cos(ka/2) = \sqrt{1-\left(\frac{\omega}{2\Omega}\right)^2} = \frac{1}{2\Omega}\sqrt{4\Omega^2-\omega^2}.$$

Thus, eqn (2.13) can be written: $g_1(\omega) = \frac{L}{a}\frac{1}{\pi}\frac{1}{\sqrt{4\Omega^2-\omega^2}} = \frac{N}{\pi}\frac{1}{\sqrt{4\Omega^2-\omega^2}}$.

Using positive k, the integral of $2g_1(\omega)$ (note that the factor of 2 is needed to account for the negative k values) over ω yields N, as expected. As mentioned in the text, writing $g_1(\omega)$ in this form offers no clear advantage.

[4] Integration in k-space yields the number of modes: $2\int_0^{\pi/a} dk\rho_1(k) = 2\frac{\pi}{a}\frac{L}{2\pi} = \frac{L}{a} = N$.

Likewise: $\int d\omega\, g(\omega) = \int d\omega \frac{L}{2\pi}\frac{1}{\Omega a\cos(ka/2)}$. Using $\omega = 2\Omega\sin(ka/2)$ again yields N.

Example 2.1. Driven Classical Oscillator with Loss

The driven 1D classical harmonic oscillator with loss is a common and useful pedagogical model that is germane to a number of areas. In the present context, the driving force might be thought of as due to the presence of electromagnetic radiation whose frequency is in near-resonance with some optical phonon frequencies. For example, suppose laser radiation is used to excite modes in an optical branch. If the radiation is of a continuous nature (for example, a continuous wave laser as opposed to a pulsed laser), information can be obtained through a frequency domain measurement such as an absorption spectrum. Alternatively, if the radiation is contained in a sufficiently short duration pulse (say, in the femtosecond regime), important dynamical processes will transpire subsequent to photoexcitation. In this latter case, measuring the system's time domain response is likely to be the preferred means of interrogation.

When an optical branch phonon is removed (deactivated), phonons can be created concomitantly in the same and other optical branches, and in the acoustic branches. The parallel to be drawn with the present example is between the decay of an excited state in the phonon case, and the dissipation that takes place in a single classical oscillator that is deactivated through a frictional force. A relationship will emerge between time and frequency domain manifestations that carry over, to a large extent, to the quantum regime. Analysis of the quantum case is more demanding mathematically, so we will not get to it here. The math is not terrible, but it is time consuming and we lack the time. However, much if not most of the important understanding can be gleaned from the classical case.

To begin, a sinusoidal driving force of frequency ω is assumed to act on a particle that would, in the absence of the frictional force, undergo simple harmonic oscillation, *i.e.*, there are no anharmonic terms in the potential energy function. However, the particle experiences a frictional force that is proportional to its speed, and therefore its equation of motion is

$$m\ddot{x} + b\dot{x} + \kappa x = Fe^{i\omega t} . \qquad (i)$$

The constant F needs to be complex if we wish to keep track of the phase of the driving force, *e.g.*, $F = F_{real}\, e^{i\varphi}$, where φ is the phase of the driving force. However, this is usually unimportant, in which case F can be taken as real. The real constant b accounts for loss in the form of friction. The simplest example of such frictional force that comes to mind is that of an object falling toward the earth through the earth's atmosphere. For example, think of a cantaloupe released from a small plane 4000 meters above the ground. It reaches a constant speed that balances the frictional force presented by the atmosphere with the gravitational force. The resulting speed is $\dot{x} = ma/b$, where a is the gravitational acceleration near the earth's surface.

Dividing eqn (*i*) by the particle mass m yields the equation that is to be solved:

$$\ddot{x} + 2\beta\dot{x} + \omega_0{}^2 x = Ae^{i\omega t} . \qquad (ii)$$

The parameter $\beta = b/2m$ is introduced to simplify the manipulations that follow, and $\omega_0{}^2 = \kappa/m$. As usual, it is understood that the real part is taken at the end.

There are two parts to the solution of eqn (*ii*), a driven part that is due solely to the right hand side, $Ae^{i\omega t}$, and a homogeneous part that satisfies eqn (*ii*) with the right hand side set equal to zero. We shall first obtain the driven solution and then add the transient part, that is, the solution of the homogeneous equation. At the end, initial conditions are used to assign values to the integration constants.

Next, $x = x_0 e^{i\omega t}$ is introduced. This is a statement that the driven solution must oscillate at the frequency of the driving force after all transients have decayed to zero. The constant x_0 is complex because in general there is a phase difference between the driving force and the system's response. Introducing $x = x_0 e^{i\omega t}$ into eqn (*ii*) yields the algebraic equation

$$\left(-\omega^2 + i2\beta\omega + \omega_0{}^2\right) x_0 e^{i\omega t} = Ae^{i\omega t} . \qquad (iii)$$

The exponential terms $e^{i\omega t}$ cancel (red lines), leaving

$$x_0 = \frac{-A}{\omega^2 - \omega_0{}^2 - i2\beta\omega} \qquad (iv)$$

$$= \frac{-A}{(\omega + \tilde{\omega} - i\beta)(\omega - \tilde{\omega} - i\beta)} , \qquad (v)$$

where $\tilde{\omega}^2 = \omega_0{}^2 - \beta^2$, which is easily verified by substitution. Let us now enlist the fact that we are interested in driving the system near one of its resonances. In this case, ω is close to either $+\tilde{\omega}$ or $-\tilde{\omega}$. Obviously it cannot be near each at the same time. It is also assumed that the frictional loss parameter β is small relative to the oscillation frequencies. In other words, we are interested in resonances that are fairly sharp. These assumptions enable us to use the approximation $\tilde{\omega} \approx \omega_0$. Choosing the $\omega \approx +\omega_0$ option, eqn (*v*) becomes

$$x_0 = -\frac{A}{2\omega_0}\frac{1}{\omega - \omega_0 - i\beta} . \qquad (vi)$$

$$= -\frac{A}{2\omega_0}\left(\frac{\omega - \omega_0}{(\omega - \omega_0)^2 + \beta^2} + i\frac{\beta}{(\omega - \omega_0)^2 + \beta^2}\right) . \qquad (vii)$$

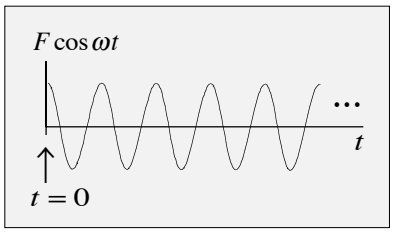
Taking the real part of $x = x_0 e^{i\omega t}$ yields the driven (steady-state) solution. The solution to the homogeneous equation (the transient solution) is added to the steady state solution to obtain the final result:

$$x = \mathrm{Re}\left(x_0 e^{i\omega t} + x_{\text{transient}}\right) \qquad (viii)$$

$$= -\frac{A}{2\omega_0}\frac{1}{(\omega-\omega_0)^2+\beta^2}\left((\omega-\omega_0)\cos\omega t-\beta\sin\omega t\right)$$

$$+\ e^{-\beta t}\left(C_1\cos\omega_0 t + C_2\sin\omega_0 t\right). \qquad (ix)$$

It is understood that the x on the left hand side of eqns (*viii*) and (*ix*) refers to $\mathrm{Re}(x)$. Hereafter, x will be used to represent $\mathrm{Re}(x)$. As mentioned earlier, the constants C_1 and C_2 are determined by the initial conditions.



Immediately following excitation with a sinusoidal driving force that remains on for a long time after it has been switched on at t = 0 (see the box on the right), there is a transient. This transient decays with an exponential lifetime of β^{-1}, leaving just the steady-state response. At exact resonance ($\omega = \omega_0$), and after any transients have decayed to zero, x is equal to $(A/2\beta\omega_0)\sin\omega_0 t$. As expected, both the displacement amplitude x and the speed $\dot{x}$ vary inversely with the frictional loss parameter β. For example, notice that as the loss goes to zero, x and $\dot{x}$ each goes to infinity. This is not surprising, however, because the model is unphysical in the sense that the displacement and speed must increase without bound for a driven system without loss. All physical systems have some degree of loss (coupling to something they perceive as exterior), so think of this divergence as just one of those things that happens from time to time with mathematical models.

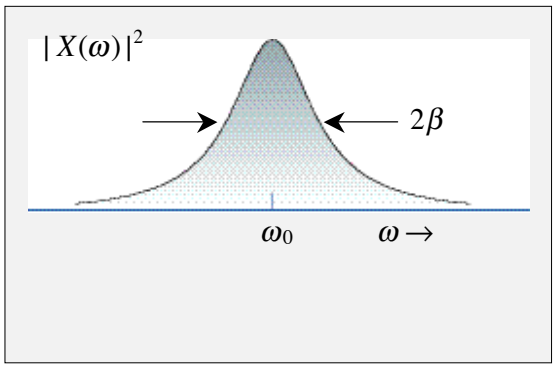
A sinusoidal driving force has been used in the above example. However, an arbitrary driving force can be expressed mathematically using a Fourier expansion. For example, think of a repetitive driving force comprised of harmonics of a driving force's fundamental frequency. Equation (*ii*) then becomes many equations, namely, one for each harmonic. The solutions to these equations follow straightforwardly. They are essentially the same as for the case of a single driving frequency, but with different frequencies.

If the driving force is absent, we are left with just the transient response. For example, with $\dot{x}$ = 0 at t = 0 we have $x(t) = C_1 e^{-\beta t}\cos\omega_0 t$. This is a rapid oscillation whose envelope is damped exponentially. To see what this looks like in the frequency domain, take the Fourier transform of $x(t)$. Suppressing all constants, this is given by

$$X(\omega) = \int_0^\infty dt\, e^{-i\omega t}e^{-\beta t}\left(e^{i\omega_0 t}+e^{-i\omega_0 t}\right) \qquad (x)$$

$$= \frac{-i}{(\omega-\omega_0)-i\beta}+\frac{-i}{(\omega+\omega_0)-i\beta}. \qquad (xi)$$

The second term is negligible as long as we restrict frequencies to $\omega \approx \omega_0$. There are several complementary ways to view this result. Perhaps the simplest is to note that its squared modulus is proportional to a Lorentzian lineshape, as shown on the right. For $\omega \approx \omega_0$, eqn (*xi*) reduces to a single complex entity (the first term on the right hand side) that can be expressed in terms of its real and imaginary parts:



$$X(\omega) = \frac{\beta^2}{(\omega-\omega_0)^2+\beta^2} - i\frac{(\omega-\omega_0)}{(\omega-\omega_0)^2+\beta^2}.$$

And so on.

2.5. Three-Dimensional Lattice

Let us now extend $g_1(\omega_k)$ from 1D to 3D. Again, we shall start with the density in *k*-space, but this time for a 3D volume element

$$\rho_3(k_x,k_y,k_z)\,d^3k = (L_xL_yL_z/(2\pi)^3)\,dk_x\,dk_y\,dk_z\,. \tag{2.14}$$

It is a safe assumption that the heat capacity of a crystal does not depend on its shape.[5] Therefore, $L_xL_yL_z = V$ is used. Now replace $dk_x\,dk_y\,dk_z$ with the spherical volume element: $k^2\sin\theta\,d\theta\,d\phi\,dk$. Integration over θ and ϕ yields 4π, which gives $\rho_3(k)\,dk = (V/2\pi^2)\,k^2dk$. Thus, for 3D, $g(\omega)$ is

$$\boxed{g(\omega) = \frac{V}{2\pi^2}k^2\frac{dk}{d\omega}\,,} \tag{2.15}$$

where $dk/d\omega$ is evaluated using the dispersion relation.

The $g(\omega)$ in eqn (2.15) is the 3D counterpart to $g_1(\omega)$ in eqns (2.11) and (2.13). It refers to a single branch of the dispersion relation. To include all participating branches, we simply add their contributions. This includes the optical branches in addition to the acoustic branches. Optical branches are important when k_BT is comparable to or exceeds the optical phonon frequencies. We shall avoid the math required to deal directly with the inclusion of the optical branches by using the Debye model, as described below.

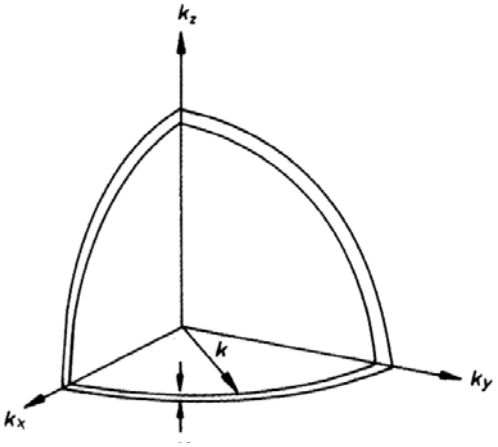


Figure 4. The differential volume element in *k*-space is a thin spherical shell of thickness *dk* (from Hook and Hall).

Referring to Fig. 4, it is understood that *k* is positive because it is the radial distance in *k*-space. As mentioned earlier, the subscript *k* on ω_k is no longer needed, as frequency and wave vector are, for all practical purposes, continuous variables. This is implicit in eqn (2.15), where we write $dk/d\omega$. The term $dk/d\omega$ is obtained from the dispersion relation. The average energy of the sample at temperature *T* is obtained by integrating the product of $g(\omega)$ given by eqn (2.15) times the average energy contained in each mode at temperature *T*. This latter quantity is ω_k times $\langle n_k \rangle$ given

[5] A crystal's thermal conductivity in general varies, often strongly, with the direction of heat flow. On the other hand, in the absence of a thermal gradient, energy density is constant to a high degree of accuracy throughout the crystal, influenced only by the crystal edges.

by eqn (2.7), though now the subscripts are dropped and $\hbar = 1$ is used. Integration is over all modes. Putting all of this together, the average stored energy is given by

$$\langle E \rangle = \int_0^\infty d\omega \left(\frac{\omega}{e^{\omega/k_B T} - 1} \right) g(\omega) \tag{2.16}$$

These are restricted to the first Brillouin zone.

$$= \frac{V}{2\pi^2} \int_0^\infty d\omega \left(\frac{\omega}{e^{\omega/k_B T} - 1} \right) k^2 \frac{dk}{d\omega} . \tag{2.17}$$

The function $g(\omega)$ is truncated at some maximum value in accord with k being restricted to the first Brillouin zone. This establishes the upper integration limit. In other words, the limit of infinity in eqn (2.16) is never reached. It is retained however (with the limit imposed by the first Brillouin zone understood), because it simplifies some of the integrations encountered below.

As mentioned above, though not indicated explicitly, the integration in eqn (2.17) is carried out over all branches. In other words, there are separate integrations for each branch because a separate $g(\omega)$ applies to each branch. If the crystal is monatomic (a single spring constant), there are three integrals, one for each of the acoustic branches. With less simple crystals, optical branches enter. In general, dispersion relations cannot be expressed in closed form, and they can be complicated. Thus, we can appreciate why an accurate calculation could prove arduous, perhaps not as a research project, but certainly as a classroom exercise.

From eqn (2.16), it is easy to see what happens at high temperature. You probably have encountered the high-T limit before, say, in undergraduate physical chemistry. The contents of the large parentheses reduces to $k_B T$, and eqn (2.16) becomes

$$\langle E \rangle = k_B T \int_0^\infty d\omega \, g(\omega) . \tag{2.18}$$

The value of the integral is, by definition, $3N$ for a monatomic 3D lattice. That is, N atoms have $3N$ degrees of freedom that result in $3N$ vibrational modes. Thus, differentiation of $\langle E \rangle$ with respect to T gives a simple expression for the high temperature limit of the heat capacity:

$$\boxed{C = 3N k_B \quad \text{high-}T \text{ limit}} . \tag{2.19}$$

This result also follows from the fact that, classically, each vibrational degree of freedom contributes k_B to the heat capacity, and there are $3N$ degrees of freedom. The heat capacity given by eqn (2.19) is twice as large as for a gas comprised of N atoms. The difference is due to the fact that vibrational energy varies quadratically with both momentum and displacement, whereas a gas has only the momentum contribution.

Turning to the low temperature limit, we now invoke the fact that the acoustic dispersion curves are linear in this regime, with speeds of sound v_{L}, v_{T_1}, and v_{T_2} for the longitudinal and two transverse branches, respectively. Using this with eqn (2.17) yields

$$\langle E \rangle = \frac{V}{2\pi^2}\underbrace{\left(\frac{1}{{v_{\mathrm{L}}}^3}+\frac{1}{{v_{\mathrm{T}_1}}^3}+\frac{1}{{v_{\mathrm{T}_2}}^3}\right)}_{\text{sum over 3 acoustic branches}}\int_0^\infty d\omega\,\frac{\omega^3}{e^{\omega/k_BT}-1}\,. \tag{2.20}$$

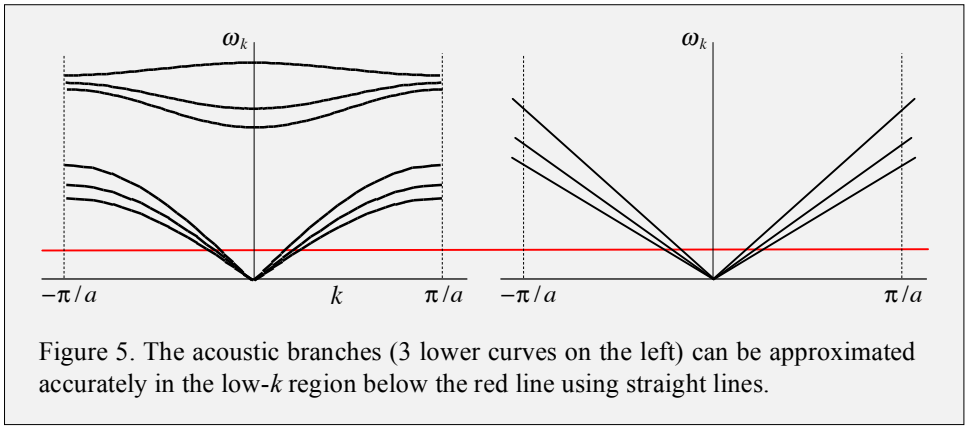


Figure 5. The acoustic branches (3 lower curves on the left) can be approximated accurately in the low-*k* region below the red line using straight lines.

Figure 5 illustrates the use of the linear portions of the three acoustic branches. The upper integration limit of infinity is acceptable because high frequency modes do not participate at low temperatures. In the spirit of keeping things simple, the large parenthetic term in eqn (2.20) is replaced by $3v^{-3}$. You can think of this as an average value of the inverse cube of the speed of sound in the acoustic branches, or, equivalently, that we are assuming that the speed of sound is the same for each of the branches. In any event, using $3v^{-3}$ and the substitution $x = \omega / k_BT$, eqn (2.20) becomes

$$\langle E \rangle = \frac{V}{2\pi^2}\frac{3}{v^3}\frac{(k_BT)^4}{\hbar^3}\int_0^\infty dx\,\frac{x^3}{e^x-1}\,. \tag{2.21}$$

Note that explicit $\hbar$ has now been introduced. The value of the integral is $\pi^4/15$, and differentiation with respect to T gives the heat capacity for the low temperature limit

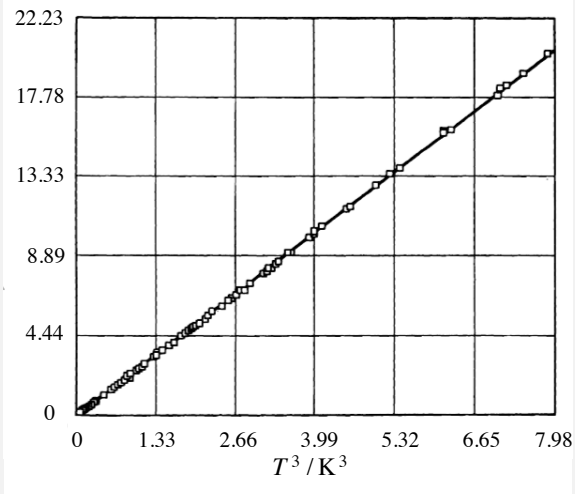
$$C = 3.95\left(\frac{k_BV}{v^3}\right)\left(\frac{k_BT}{\hbar}\right)^3 \quad \text{low-}T\text{ limit}\,. \tag{2.22}$$

This model of the low temperature limit holds up well against experimental data. Figure 6 shows the result obtained by carrying out heat capacity measurements for a

monatomic lattice of argon atoms. The temperature range is low indeed, not exceeding 2 K. The agreement between experiment and theory is stunning.

The limiting cases discussed above are straightforward. Let us now examine the full temperature range by using the model introduced by Debye. This model has great value because it enables one to judge material properties on the basis of parameters that can be obtained from experiment. It is an interpolation that gets around the mathematical task that would be faced were we to deal directly with the optical modes.[6]

Figure 6. Heat capacity in units of $\mathrm{mJ \cdot mol^{-1} \cdot K^{-1}}$: experimental result for a monatomic lattice of solid argon at temperatures < 2 K. The data follow the predicted T^3 dependence very well (from Kittel, *Introduction to Solid State Physics* [5], who references L. Finegold and N. E. Phillips, for the data shown on the right).



Keep in mind that dispersion relations in general can be complicated, depending on what goes into a primitive unit cell, crystal symmetry, and so on. Figure 7 shows the case of a diatomic lattice of a typical face-centered cubic crystal. This is still relatively simple, yet we see changes from the dispersion relations introduced so far. For example, one optical branch increases with wave vector while the other decreases with wave vector, and two of the LA acoustic branches are of different character than those introduced so far. Notice that some of the curves cross in *k*-space.

[6] A convenient and accurate algorithm for dealing with many vibrational degrees of freedom is the one introduced by Beyer and Swinehart [75]. It is used frequently to estimate level densities in statistical theories of unimolecular reactions.

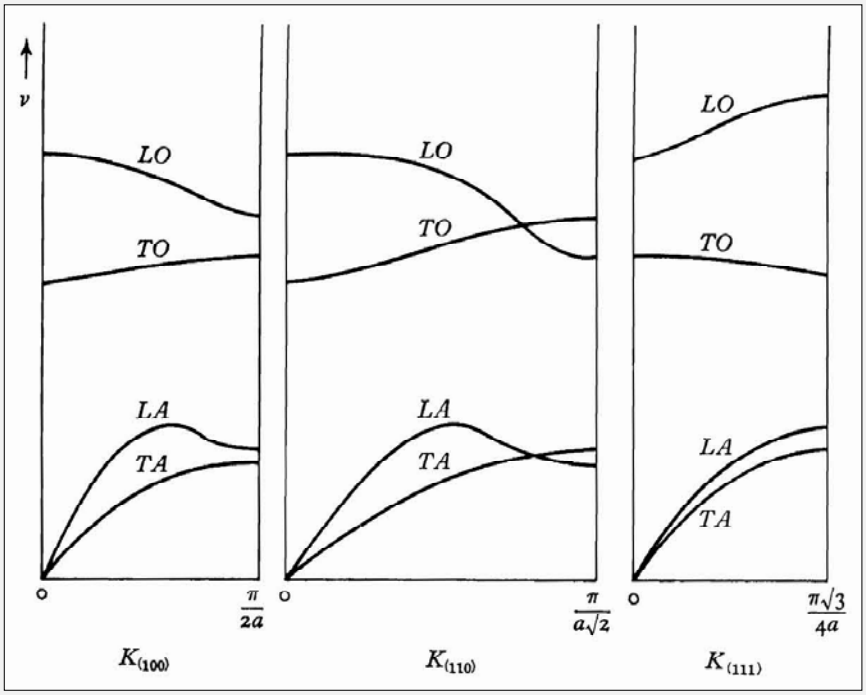


Figure 7. From Sherwood: *Vibrational Spectroscopy of Solids*. Note that for the (100), (110), and (111) directions, the transverse phonons are doubly degenerate.

2.6. Debye Model


Peter Debye

In the simplest version of the Debye model, the speed of sound is assigned the same value in each of three acoustic branches, and the dispersion relation is taken to be a straight line: $\omega = vk$. There is no flattening of the $\omega(k)$ curve as k approaches the edge of the Brillouin zone. The optical branches are not included explicitly. Their effect insofar as heat capacity is subsumed into the dispersion relation $\omega = vk$.

The relation $\omega = vk$ is taken as valid up to some cutoff frequency, ω_D, that is referred to as the Debye frequency, above which there are no available modes. The cutoff wave vector associated with ω_D is $k_D = \omega_D / v$. We shall see that k_D lies within the Brillouin zone, not at its edge. This construct might sound bizarre, because you have learned that dispersion curves, even for the acoustic branches, are never so simple. Or are they?

The flattening of the $\omega(k)$ curve as k approaches the edge of the Brillouin zone is a consequence of the discrete nature of the lattice. Were the medium completely elastic with harmonic restoration, and with no hint of an underlying lattice structure, there could be neither the above flattening nor optical branches. We shall return to this situation later and see that it represents some very important systems. For the case at hand, the Debye ansatz ensures a model that is exact in the high and low frequency limits. In other words, it is an interpolation between these limiting cases. Therefore it might have problems in the intermediate regime. Nonetheless, it serves as a reasonable first estimate, or at least a pedagogical exercise on the way to more advanced models.

The cutoff frequency ω_D is obtained by using the fact that the integral of $g(\omega)d\omega$ from 0 to ω_D, by definition, gives the number of modes in 3D space. Equivalently, integration of the k-space density from $k = 0$ to k_D must also give the number of modes. We shall now use $g(\omega)$ to obtain ω_D, in which case integration is carried out in frequency space. For N atoms, there are $3N$ modes. Introducing the above assumptions into eqn (2.15) and integrating $g(\omega)d\omega$ yields:

$$3N = \int_0^{\omega_D} d\omega\, g(\omega) \tag{2.23}$$

Using $g(\omega) = \rho(k)(dk/d\omega)$ and eqn (2.15) gives

$$3N = \frac{V}{2\pi^2} \int_0^{\omega_D} d\omega \underbrace{3k^2 \frac{dk}{d\omega}}_{3(\omega^2/v^3)} \qquad \text{3 acoustic branches} \tag{2.24}$$

$$= \frac{V}{2\pi^2}\frac{3}{v^3}\int_0^{\omega_D} d\omega\, \omega^2 \tag{2.25}$$

Thus, $\omega_D{}^3 = v^3 6\pi^2 N/V$ and $k_D{}^3 = 6\pi^2 N/V$. The parameters ω_D and k_D are seen to have quite simple forms:

$$\boxed{\omega_D = v k_D = v(6\pi^2 n)^{1/3},} \tag{2.26}$$

where $n = N/V$ is the number of atoms per unit volume.

An alternate way to arrive at this result is to use the volume of a *k*-space sphere having a radius of k_D, namely $4\pi k_D{}^3/3$. It is assumed that all modes are contained within this sphere. Namely, they fill *k*-space up to the sphere's edge but not beyond. Dividing the volume $4\pi k_D{}^3/3$ by the *k*-space volume of a single *k*-space cell: $(2\pi)^3/V$, gives the number of modes per branch. Multiplying this number by 3 accounts for the 3 branches.

Combining the above items and using the fact that there are $3N$ modes gives

$$3N = \left(\frac{4\pi}{3}k_D{}^3\right)\left(\frac{3V}{(2\pi)^3}\right) = \frac{3V}{6\pi^2}k_D{}^3 \tag{2.27}$$

Thus, $k_D = (6\pi^2 n)^{1/3}$, which is eqn (2.26). The only free parameter in the model is the speed of sound – equivalently the Debye frequency: $\omega_D = v(6\pi^2 n)^{1/3}$. This can be taken from independent measurements or adjusted to fit data. However, it should not deviate much from the known speed of sound in the material.

Now that we have the Debye cutoff frequency ω_D, the heat capacity is obtained in a few steps starting with eqn (2.17). Note that the right hand side of eqn (2.17) needs to be multiplied by 3 to account for the 3 acoustic branches. With the upper integration limit set at ω_D, eqn (2.17) becomes

$$\langle E \rangle = \frac{3V}{2\pi^2}\int_0^{\omega_D} d\omega \left(\frac{\omega}{e^{\omega/k_BT}-1}\right) k(\omega)^2 \frac{dk(\omega)}{d\omega}. \tag{2.28}$$

With the Debye ansatz, $\omega = vk$, the heat capacity is given by

$$C = \frac{3V}{2\pi^2 v^3}\frac{\partial}{\partial T}\int_0^{\omega_D} d\omega \left(\frac{\omega^3}{e^{\omega/k_BT}-1}\right). \tag{2.29}$$

To evaluate the right hand side, take the derivative and carry out the integration. The steps are given in the box below. Also, let us express the heat capacity as a per unit volume quantity by dividing the right hand side of eqn (2.29) by the volume *V*. The same symbol *C* is used, hopefully without confusion. The resulting expression is

$$C = 9nk_B\left(\frac{T}{\Theta_D}\right)^3 \int_0^{x_D} dx\left(\frac{x^4 e^x}{(e^x-1)^2}\right), \tag{2.30}$$

where $\Theta_D = \omega_D / k_B$ and $x = \omega / k_B T$.

This result is general in the sense that it is valid for any temperature. The parameter Θ_D can be obtained from experimental measurements. It is used frequently as a rough measure of the stiffness of a crystal. Values for alkali halide ionic crystals are given in Table 1, and values for crystals made of atoms are given in Table 2.

Using $\omega_D{}^3 = 6\pi^2 v^3 (N/V)$, eqn (2.29) becomes: $C = \frac{9N}{\omega_D{}^3}\int_0^{\omega_D} d\omega\,\omega^3 \frac{\partial}{\partial T}\left(\frac{1}{e^{\omega/k_BT}-1}\right)$

Next, carry out a couple of operations:

$$C = \frac{9N}{\omega_D{}^3}\int_0^{\omega_D} d\omega\,\omega^3\left(\frac{(\omega/k_BT^2)e^{\omega/k_BT}}{(e^{\omega/k_BT}-1)^2}\right) = \frac{9N}{\omega_D{}^3}\frac{1}{k_BT^2}\int_0^{\omega_D} d\omega\left(\frac{\omega^4 e^{\omega/k_BT}}{(e^{\omega/k_BT}-1)^2}\right)$$

With $x = \omega / k_B T$, this becomes:

$$C = \frac{9N}{\omega_D{}^3}\frac{(k_BT)^5}{k_BT^2}\int_0^{x_D} dx\left(\frac{x^4 e^x}{(e^x-1)^2}\right) = 9Nk_B\left(\frac{k_BT}{\omega_D}\right)^3 \int_0^{x_D} dx\left(\frac{x^4 e^x}{(e^x-1)^2}\right)$$

Using the definition of the Debye temperature: $\Theta_D = \omega_D / k_B$, the above expression becomes:

$$C = 9Nk_B\left(\frac{T}{\Theta_D}\right)^3 \int_0^{x_D} dx\left(\frac{x^4 e^x}{(e^x-1)^2}\right) \xrightarrow{\div V} C = 9nk_B\left(\frac{T}{\Theta_D}\right)^3 \int_0^{x_D} dx\left(\frac{x^4 e^x}{(e^x-1)^2}\right)$$

In the high temperature limit the integral is evaluated by noting that the exponential in the numerator is equal to one, while the denominator is equal to x^2 because $e^x \to 1 + x$. The integrand is therefore x^2, and integration yields $x^3/3$. Thus, we see that the high temperature limit of eqn (2.30) is $C = 3nk_B$, which is recognized as the classical limit of the heat capacity for a crystal comprised of N particles (*i.e.*, $3N$ oscillators) in a volume V. Recall eqn (2.19).

The low temperature limit also follows straightforwardly from eqn (2.30). In this case, the upper integration limit is set at infinity for convenience. This works because the occupied modes are nearly all contained in the low-k region at low temperature, so the usual limit of restriction to the first Brillouin zone can be relaxed, and k can be allowed to go to infinity. The integral in eqn (2.30) is one of those things that happen every so often. I am not going to discuss its evaluation. The final result is that the heat capacity is

Table 1. Debye temperatures (Θ_D) for alkali halide crystals

	F	Cl	Br	I
Li	730	422	—	—
Na	492	321	224	164
K	336	231	173	131
Rb	—	165	131	103

$$C = 234\, n k_B \left(\frac{T}{\Theta_D}\right)^3 \quad \text{low-}T\text{ limit} \tag{2.31}$$

Figure 8 shows a plot of the general expression for the heat capacity according to the Debye model [the left entry is a plot of eqn (2.30)] and experimental results for Si and Ge. Agreement between experiment and theory is very good. You can see this for germanium because I altered the horizontal scale in the figure I took from Kittel. Though less obvious, silicon also agrees, albeit over a smaller range.

Table 2. Debye temperatures (Θ_D) for elements

Li	400	Be	1000	B	1250	C	1860	As	285
Na	150	Mg	318	Al	394	Si	625	Sb	200
K	100	Ca	230	Ga	240	Ge	360	Bi	120
				In	129	Sn	260		
				Tl	96	Pb	88		

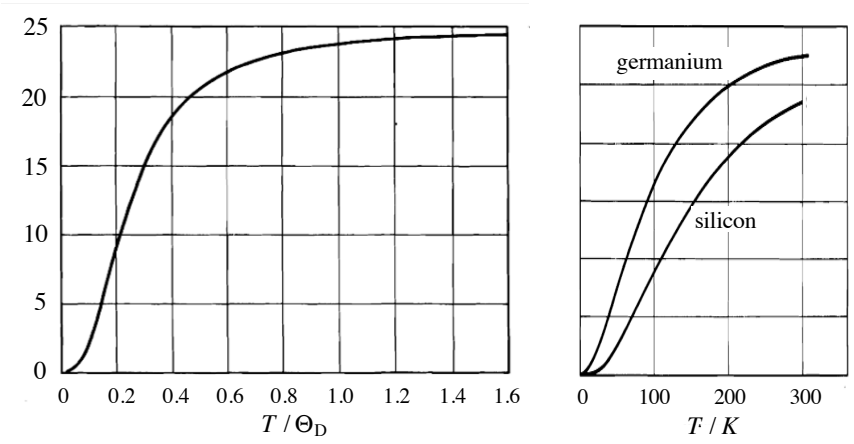


Figure 8. Left: calculated heat capacity versus T/Θ_D using eqn (2.31). Right: data for silicon and germanium follow Debye-like curves, assuming appropriate scale changes for horizontal axis (adapted from Kittel).

2.7. Thermal Conductivity

The remainder of Chapter 2 deals with thermal conductivity. This might sound like a boring subject, a leftover from the dark ages of classical thermodynamics. In fact, it is both scientifically challenging and of enormous technological importance. The number of areas in which thermal conductivity plays an important role is huge. We cannot begin to cover them here. In the next few sections, several fundamental aspects that are germane to a wide range of crystalline media will be introduced and examined.

We shall begin by considering the fact that the thermal conductivity of a perfectly harmonic crystal is infinite. Phonon wave packets can flow throughout such a crystal without changing their composition. They might broaden, distort, etc., but they remain composed of the same distribution of k values that they began with. There is no mechanism for excitation to be transferred among the various normal modes. Consequently, heat can flow in the absence of a temperature gradient, which does not happen except in superfluid media, well known examples around here being superconductivity and liquid helium nanodroplets, which cool spontaneously to ~ 0.4 K.

Aside from scattering by defects, impurities, and surfaces, the only mechanism for wave vector change is anharmonicity, which, by definition, is absent in a perfectly harmonic crystal. The anharmonic parts of the potential that are important are those that couple the normal modes to one another. This is akin to the well-known couplings that exist in molecules, such as Fermi resonances, Darling-Dennison resonances, and so on. In dealing with anharmonicity in crystalline lattices, it will be seen that things can get complicated but at the same time quite interesting.

In Section 2.8 (immediately below), a simple expression will be derived for the thermal conductivity of a crystal in terms of heat capacity C, the speed of sound v, and a scattering parameter – either a scattering length l or, equivalently, a scattering time τ, where $l = v\tau$. We will find that a simple form is obtained, namely, $\kappa = Cvl/3$, where κ is the thermal conductivity. The model is crude, but it contains sensible ingredients. The trickiest one is the scattering length l. Most of our time will be spent discussing this, finishing with the high and low temperature limits. As an interlude, we will explore an interesting phenomenon due to anharmonicity in a gas of carbon monoxide molecules. Time permitting this will be presented in one of our discussion sessions, so it is only outlined in these notes.

Material	κ $(\mathrm{W \cdot m^{-1} \cdot K^{-1}})$
air	0.025
water (liquid)	0.6
water (ice)	2
concrete	1.7
glass	1.1
gold	318
lead	35
copper	401
silver	429
diamond	900-2320

Thermal conductivity values: The large range for diamond is due to anisotropy.

2.8. Model Adapted from Gas Phase Kinetics

To begin, consider the energy per unit area per second that passes through a surface, as indicated in Fig. 9. It is assumed, with no loss of generality, that the surface normal lies parallel to the direction of heat flow. We seek an expression for the net flux in the direction of interest, namely, in the direction of the temperature gradient.

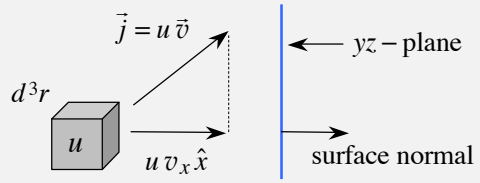


Figure 9. The energy contained in a small volume is $u\,d^3r$. It passes through the yz-plane (blue line). The flux $\vec{j}$ (whose units are $J \cdot m^{-2} \cdot s^{-1}$) is equal to $u\vec{v}$. Its component in the direction of the surface normal is $u\,v_x\hat{x}$.

The strategy that underlies this approach derives from gas phase kinetics. Referring to Fig. 9, u denotes the energy density and $\vec{j}$ its corresponding flux, *i.e.*, $\vec{j} = u\vec{v}$. A small volume d^3r containing the energy density u is some distance from a surface that is nothing more than the yz-plane (solid blue line in Fig. 9, and dashed blue line in Fig. 10). The net flow of energy through this surface from the higher-temperature region to the lower-temperature region is the desired flux. It is assumed that temperature changes smoothly and monotonically in going from higher to lower temperature.

Scattering Ansatz

Referring to Fig. 10, assume that the energy density u is contained in a small volume that travels a length l without scattering. This is akin to elementary gas phase kinetics theory, where a mean distance between collisions is invoked. Because the gradient of the energy density is in the x-direction, we write $u(x)$ to denote energy density as a function of distance along the x-axis. It is assumed that $u(x)$ does not vary in the yz-plane.

Figure 10. It is assumed that phonon wave packets travel, on average, a distance l before making a "collision." Those that arrive at the black dot on the yz-plane (blue dashed line) originated on a sphere of radius l. The system has azimuthal symmetry about the x-axis. There are more phonons on the left than on the right, as these are the higher and lower-T regimes, respectively. Thus, there is net flow of energy to the right.

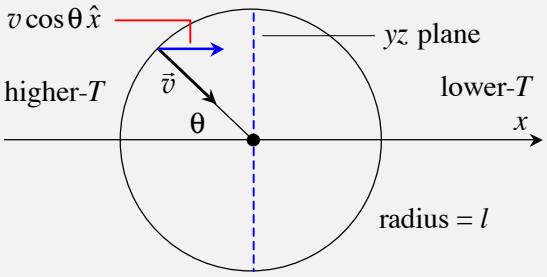


Figure 10 is a clever means of visualizing the flux that passes through a surface whose normal lies along the temperature gradient. To see how this works, consider a point in the yz-plane. This point is taken as the origin in the figure, and we ask the question: If energy

density moves a distance l at speed v, where must it be located, and in what direction must it be traveling, in order that it reaches the origin after traveling the distance l?

Clearly, it must be distributed on a sphere of radius l with velocities that point at the origin. This flow of energy contributes to the flux in the x-direction in proportion to the x-component of the velocity, $v_x = v\cos\theta$. Keep in mind that the *net* flux that passes through the blue plane is the difference between energy that passes from left to right (originates from the left hand side of the sphere) and energy that passes from right to left (originates from the right hand side of the sphere). The x-directed flux will be evaluated at the origin with the understanding that the result applies anywhere in the yz-plane. The net x-directed flux through the surface is obtained by averaging the product $v_x u(x) = v\cos\theta u(x)$ over the surface of the sphere:

$$j_x = \frac{1}{4\pi}\int_0^{2\pi} d\phi \int_0^{\pi} d\theta \sin\theta \; v\cos\theta u(x) . \tag{2.32}$$

Nothing depends on ϕ, so integration over ϕ gives 2π.

Next, to introduce the spatial variation of $u(x)$ in the direction of the temperature gradient, $u(x)$ is expanded in a Taylor series: $u(x) = u(0) + (\partial u/\partial x)_0 x \ldots$. The constant term in the expansion, $u(0)$, does not contribute to j_x because there can be no net flux through the surface if the energy density is constant throughout space. The first non-vanishing term is what we will use, and this yields, using the substitution $x = -l\cos\theta$ (see Fig. 10):

$$j_x = \tfrac{1}{2} v \int_0^{\pi} d\theta \sin\theta\cos\theta \left(\left.\frac{\partial u}{\partial x}\right|_{x=0} \left(-l\cos\theta\right)\right), \tag{2.33}$$

where $(\partial u/\partial x)\big|_{x=0}$ is evaluated at the $x = 0$ location indicated by the origin in Fig. 10. It is implicit in the above that l is smaller than the characteristic distance over which $u(x)$ changes significantly. Hereafter, the subscript on the term $(\partial u/\partial x)\big|_{x=0}$ shall be taken as understood. Thus, the expression for j_x in eqn (2.33) becomes

$$j_x = -\tfrac{1}{2} v l \frac{\partial u}{\partial x} \int_0^{\pi} d\theta \sin\theta\cos^2\theta . \tag{2.34}$$

The integral is evaluated by using $d\theta\sin\theta\cos^2\theta = -d(\cos\theta)\cos^2\theta = -d\alpha\,\alpha^2$. Integration yields $2/3$. The term $\partial u/\partial x$ is equal to $(\partial u/\partial T)(\partial T/\partial x) = C(\partial T/\partial x)$. Thus, the expression for the x-directed flux becomes

$$j_x = \tfrac{1}{3} C v l \left(-\frac{\partial T}{\partial x}\right), \tag{2.35}$$

or

$$j_x = -\kappa \frac{\partial T}{\partial x}, \tag{2.36}$$

where κ is the thermal conductivity:

$$\boxed{\kappa = \tfrac{1}{3} C v l .} \tag{2.37}$$

The C in eqn (2.37) is the heat capacity. Its units are $\mathrm{J \cdot m^{-3} \cdot K^{-1}}$. Thus, thermal conductivity κ has units $\mathrm{J \cdot m^{-1} \cdot s^{-1} \cdot K^{-1}}$ (equivalently $\mathrm{W \cdot m^{-1} \cdot K^{-1}}$), and the x-directed flux j_x has units $\mathrm{J \cdot m^{-2} \cdot s^{-1}}$ (equivalently $\mathrm{W \cdot m^{-2}}$). As mentioned earlier, the tricky part is the scattering length. Let us now address this starting with crystal momentum.

2.9. Phonon Crystal Momentum

To understand how thermal conductivity works in an electrically insulating crystal, it is necessary to understand how "crystal momentum" (wave vector) transfer works in the crystal. We have seen that the concept of momentum in a periodic lattice can be subtle, so let us begin with this. In the exercise carried out below, the momentum of the overall 1D lattice will be calculated using classical mechanics. Following this, the analogous quantum mechanical calculation will be carried out.

To lessen confusion over nomenclature, the label P is assigned to the crystal's classical momentum. For a 1D monatomic lattice, P is defined as the mass of a single particle times the sum of the time derivatives of the individual displacements. It is important to distinguish the momentum of the crystal as a whole (the motion *of* its center-of-mass (c.m.) relative to a lab reference frame) from the motions that take place inside the crystal, that is, without changing the crystal's overall momentum relative to a lab reference frame. The analogous situation in gas phase dynamics is the separation of a system into motion *of* a c.m. system and motions that transpire *within* the c.m. system.

In the analogous quantum mechanical case, the normal momentum $\hat{P}_k$ shall be discussed; specifically, its expectation value for the k^{th} normal mode will be calculated. We will see that this also returns a value of zero. In both the classical and quantum cases, we leave aside for the time being $k = 0$.

Classical Case

With sinusoidal time variation $e^{i\omega_k t}$ or $e^{-i\omega_k t}$ assigned to all displacements associated with the k^{th} normal mode (either exponential can be chosen), the complex displacement at the n^{th} site is given by $A_n = A e^{ikan} e^{\pm i\omega_k t}$, where A is real. Again, keep in mind that all physical quantities are obtained by taking real parts. Differentiation with respect to time results in multiplication by $\pm i\omega_k$. Thus, using the above complex exponentials to express the time behavior, the classical momentum P is given by

$$P = m\sum_n \frac{dA_n}{dt} \tag{2.38}$$

$$= m(\pm i\omega_k)A \underbrace{\left\{\sum_n e^{ikan}\right\}}_{=0} e^{\pm i\omega_k t}. \tag{2.39}$$

Site-to-site phase variation of the displacements is accounted for entirely with the phase factors e^{ikan}. We saw previously that the sum inside the bracket in eqn (2.39) vanishes identically. Thus, the step in which the real part of P is taken never arises because P is equal to zero identically. In other words, an individual phonon normal mode (single k value) has no net momentum. It is assumed of course that no external force acts on the crystal to cause it to move as a whole. You might find this surprising because we saw earlier that k is associated with a direction and amount of phase progression.

The overall momentum P is composed of contributions from each of the n sites. It is the site-to-site phase progression that results in the overall momentum being zero. The mathematical statement of this is that $\sum_n e^{ikan} = 0$. What is going on physically is not difficult to visualize, for example, with the diagrams in Chapter 1, Fig. 7 and the transverse displacements in Chapter 1, Fig. 16. For any given k value, half of the momenta point in one direction, while the other half point in the opposite direction. Therefore the net momentum is zero for each k value. The same idea applies to longitudinal compression-rarefaction waves, but these do not make as nice a picture. Thus, as long as $k \neq 0$, there is no overall momentum. Moreover, recall that $k = 0$ is not a vibrational mode. It is overall translation.

Quantum Case

Let us now return to the relationship between P and $\hat{P}_k$. In the above, P was treated classically, and it was shown that $P = 0$ for a normal mode. This does not mean that the classical momentum P vanishes everywhere. Obviously this cannot be so, or there would be no motion whatsoever, and the system would remain at 0 K. The phase progression e^{ikan} requires that the momentum at any given site is equal and opposite that of a counterpart momentum at another site. In other words, the phases are such that the overall classical momentum for a given phase progression (nonzero k value, which includes all of the sites) vanishes.[7]

Now consider the quantum case. Were we to calculate the quantum mechanical expectation value of $\hat{P}_k$, we would find that it also vanishes. This follows immediately (and without the need to actually carry out a calculation) from the fact that $\hat{P}_k$ acting on the k^{th}

[7] Recall the quantum mechanical harmonic oscillator for a single particle. Symmetry requires that the expectation values $\langle x \rangle$ and $\langle p \rangle$ each vanish identically, whereas the quadratic quantities do not vanish: $\langle kx^2/2 \rangle = \langle p^2/2m \rangle = (\omega/2)(v + 1/2)$. Wave packets can be used to describe back and forth oscillation within the well.

mode, $|\, n_k \,\rangle$, raises and lowers it to $|\, n_k + 1 \rangle$ and $|\, n_k - 1 \rangle$, causing the expectation value $\langle n_k \,|\, \hat{P}_k \,|\, n_k \,\rangle$ to vanish. However, $\hat{P}_k$ enters the Hamiltonian quadratically, so there is no problem with the fact that its expectation value vanishes. The expectation value of $\hat{P}_k{}^2$ does not vanish. In mathematical terms, this is seen by writing $\hat{P}_k{}^2 = \hat{P}_k \hat{P}_k$ and inserting closure between the two $\hat{P}_k$'s:

$$\langle n_k \,|\, \hat{P}_k{}^2 \,|\, n_k \,\rangle = \sum_{n_{k'}} \langle n_k \,|\, \hat{P}_k \,|\, n_{k'} \rangle \langle n_{k'} \,|\, \hat{P}_k \,|\, n_k \,\rangle \tag{2.40}$$

$$= K \sum_{n_{k'}} \langle n_k \,|\, \left(\hat{a}_k{}^\dagger - \hat{a}_{-k} \right) |\, n_{k'} \rangle \langle n_{k'} \,|\, \left(\hat{a}_k{}^\dagger - \hat{a}_{-k} \right) |\, n_k \,\rangle \tag{2.41}$$

where K is a constant. This gives the expected result. Details are assigned as an exercise.

When the lattice is not monatomic, the overall momentum still vanishes (classically and quantum mechanically). Of course, we know from common experience that lattice vibrations are transported through a crystalline lattice. Mathematically, wave packets are used to represent phonon waves traveling in any direction, dispersing, and so on. Wave packets are a more intuitive way to think of phonons than trying to imagine how waves having single k values – the normal modes – behave. Nonetheless, if all that takes place is propagation, in a perfectly harmonic crystal, of wave packets that are comprised of different numbers of quanta of the different normal modes, there is no mechanism for crystal momentum to be created or destroyed.

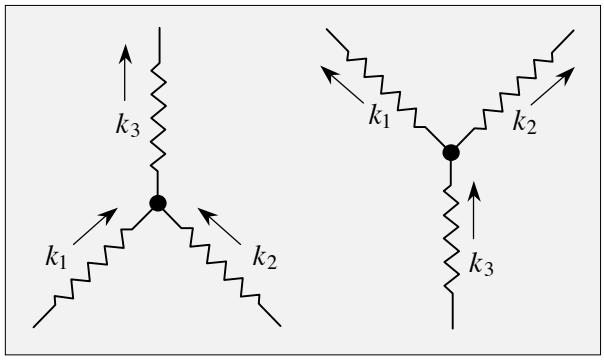
Anharmonicity

The phonons under consideration are quanta of a Hamiltonian in which the potential energy, which can be expressed in terms of a Taylor series expansion about the equilibrium position, is approximated by just the quadratic term. Obviously this is only a zero-order approximation to any realistic crystal, or for that matter any other related physical system. No potential has ever been found in a solid, liquid, or isolated molecule that contains just the quadratic term and no higher-order terms.

A Taylor series expansion of a given potential energy contains numerous higher-order terms. For a single oscillator, the potential, expanded about its ($x = 0$) equilibrium position, can be expressed as:

$$V(x) = \tfrac{1}{2}\kappa x^2 + ax^3 + bx^4 + cx^5 + \ldots.$$

For an ensemble of different modes, the number of cross terms is large, with level proximities influencing strongly the respective coupling efficiencies. All of these additional terms in a potential fall under the heading of anharmonicity. They are not simply small corrections. They are responsible for many important things. For example, without anharmonicity: (1) a solid would not undergo thermal expansion because $\langle x^2 \rangle$ would remain at its equilibrium value regardless of the degree of vibrational excitation; (2) phonons would have an infinite lifetime; and (3) most importantly for the present discussion, the thermal conductivity of a crystal would be infinite.



To see why this latter point is so, consider couplings among phonon states. Anharmonic couplings can be expressed in terms of raising and lowering operators, which makes clear their role in coupling modes to one another. As mentioned earlier, the anharmonic terms of most importance are those that couple the modes to one another. For example, suppose one quantum of one mode has energy that is approximately equal to two quanta of another mode and there is a term in the potential that couples these vibrations. This provides a pathway for deactivation and/or excitation transfer that does not exist in a purely harmonic crystal. It can be said that the phonon has a finite lifetime.

2.10. Bragg Scattering

Consider a wave incident on a crystal at some angle θ, as indicated in Fig. 11(a). It is assumed that the crystal is rigid. We shall take as our wave an *x*-ray photon with well defined energy. A slow, monoenergetic neutron can also be used. The two incoming rays are assigned equal phases in the plane indicated with a solid red line perpendicular to the rays on the left side. A phase difference accrues in the outgoing rays because the distances traveled by the rays, starting say from the red reference plane on the left, differ by $2a\sin\theta$ when the scattered rays emerge as outgoing rays. In other words, the distance between the two red planes on the right side of Fig. 3.11(a) is $2a\sin\theta$. This is also indicated in Fig. 11(b). Constructive interference occurs when $2a\sin\theta$ is equal to an integral number of *x*-ray wavelengths. This is referred to as the Bragg condition:

$$2a\sin\theta = n\lambda\,. \tag{2.42}$$

Rays are used as a means of registering phase difference. Of course, we are dealing with electromagnetic radiation, which comes in waves, not rays. For example, think of a uniform plane wave with wave vector $\boldsymbol{k}_{in}$ and uniform intensity in the plane whose normal is in the same direction as $\boldsymbol{k}_{in}$. Though it is obviously a wave, we can still use the ray picture to obtain the Bragg condition.

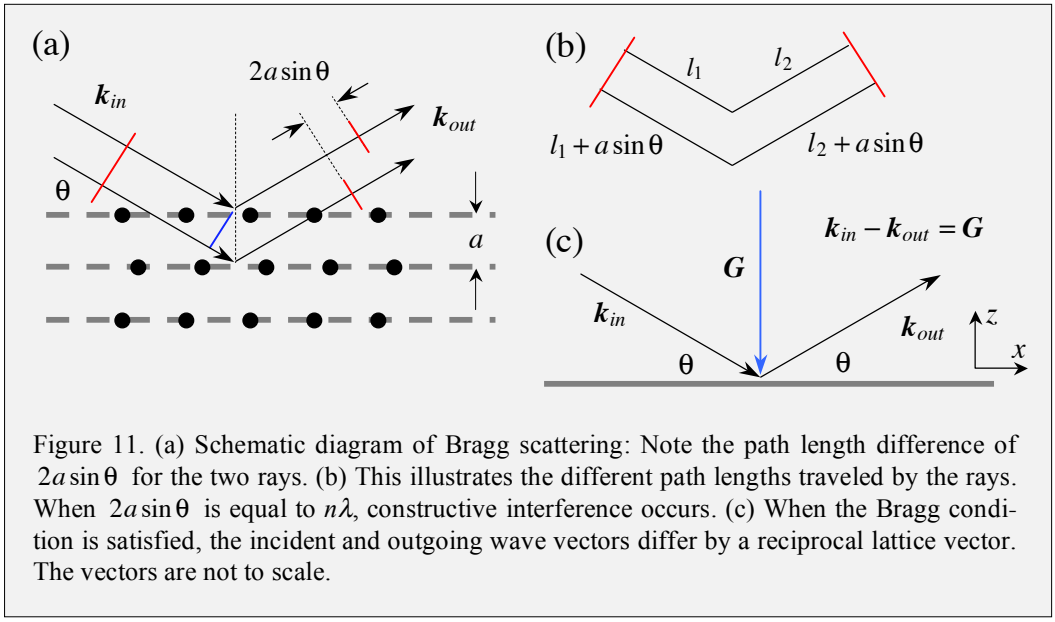


Figure 11. (a) Schematic diagram of Bragg scattering: Note the path length difference of $2a\sin\theta$ for the two rays. (b) This illustrates the different path lengths traveled by the rays. When $2a\sin\theta$ is equal to $n\lambda$, constructive interference occurs. (c) When the Bragg condition is satisfied, the incident and outgoing wave vectors differ by a reciprocal lattice vector. The vectors are not to scale.

Figure 11(c) indicates that when the Bragg condition is satisfied, $\boldsymbol{k}_{in} - \boldsymbol{k}_{out}$ is equal to a reciprocal lattice vector $\boldsymbol{G}$. For example,

$$\boldsymbol{k}_{out} = k_x \hat{x} + k_z \hat{z} = k\cos\theta\,\hat{x} + k\sin\theta\,\hat{z} \tag{2.43}$$

$$\boldsymbol{k}_{in} = k\cos\theta\,\hat{x} - k\sin\theta\,\hat{z}\,. \tag{2.44}$$

Therefore,

$$\boldsymbol{k}_{in} - \boldsymbol{k}_{out} = -2k\sin\theta\,\hat{z} \tag{2.45}$$

$$= -2\frac{2\pi}{\lambda}\sin\theta\,\hat{z} \tag{2.46}$$

$$= -2\pi\frac{2\sin\theta}{\lambda}\hat{z}\,. \tag{2.47}$$

Using the Bragg condition given by eqn (2.42) to identify the angles at which the maximum scattering intensities occur ($2a\sin\theta = n\lambda$), this becomes

$$\boldsymbol{k}_{in} - \boldsymbol{k}_{out} = -\frac{2\pi}{a}n\,\hat{z} \tag{2.48}$$

$$= \boldsymbol{G}. \tag{2.49}$$

The minus sign on the right hand side of eqn (2.48) is in accord with the direction of $\boldsymbol{G}$ and the definition of the x and z directions indicated in Fig. 11(c). You can see visually that the larger the angle θ, the larger the reciprocal lattice vector $\boldsymbol{G}$, and the larger the value of n in eqn (2.48).

Many lattice sites are involved, so a momentum change of $\hbar\boldsymbol{G}$ does not affect noticeably the movement of the lattice, and scattering can be taken to be elastic. Maxima in the scattered signal appear at θ values that satisfy the Bragg condition given by eqn (2.42). These directions (θ values) follow the rule that the difference in incident and scattered wave vectors is a reciprocal lattice vector:

$$\boxed{\boldsymbol{k}_{in} - \boldsymbol{k}_{out} = \boldsymbol{G}} \tag{2.50}$$

where $|\boldsymbol{G}| = (2\pi/a)n$. This admittedly rough description does an adequate job of describing x-ray and slow neutron diffraction / interference patterns.

2.11. A Phonon is Created

Next, consider a scattering process in which a phonon is created, as indicated in Fig. 12(a). This differs from the case presented above, where the lattice was taken as rigid. The incident and scattered wave vectors might differ by an amount that is too large for the phonon to be accommodated in the first Brillouin zone. However, we know that all phonon wave vectors must lie within the first Brillouin zone, and that a phonon wave vector that, at first sight, appears to lie outside the first Brillouin zone is equivalent to one that lies in the first Brillouin zone. The latter is obtained by compensating with whatever reciprocal lattice vector is required to get the phonon into the first Brillouin zone. In other words, the phonon wave vector, $\boldsymbol{k}$, must satisfy the relation

$$\boldsymbol{k}_{in} = \boldsymbol{k}_{out} + \boldsymbol{k}_{phonon} + \boldsymbol{G} \,. \tag{2.51}$$

Again, it is understood that $\boldsymbol{k}_{phonon}$ lies in the first Brillouin zone, and $\boldsymbol{G}$ is whatever reciprocal lattice vector is required to ensure that $\boldsymbol{k}$ lies in the first Brillouin zone. Keep in mind that eqn (2.51) refers to Fig. 12(a) rather than Fig. 11.

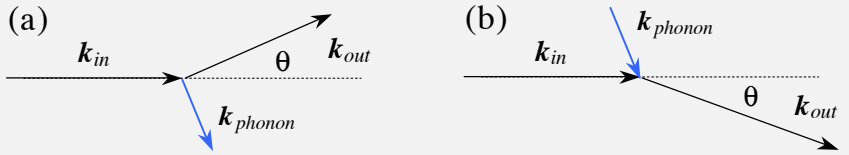


Figure 12. The phonon wave vector either (a) subtracts from, or (b) adds to the incident wave vector ($\boldsymbol{k}_{in}$) to yield the outgoing wave vector ($\boldsymbol{k}_{out}$), depending on whether the phonon is created or annihilated, respectively. Namely, in (a) a phonon is created, while in (b) a phonon is annihilated. Keep in mind that the phonon wave vector might need to be augmented by a reciprocal lattice vector to keep the phonon wave vector in the first Brillouin zone.

At the same time, conservation of momentum tells us that perhaps it is unwise to ignore the reciprocal lattice vectors $\boldsymbol{G}$ as if they did not exist, even though they have no bearing on the resulting phonon. What happens is that the crystal acquires the momentum associated with $\boldsymbol{G}$. However, because the crystal is heavy, the amount of energy imparted to it is minuscule. The smallness of the imparted energy is dramatic. The translational energy acquired by the crystal is far too small to worry about. Go ahead and make an estimate of $(\hbar G)^2/2m$ – you will see that it is tiny. In many books, you will find eqn (2.51) written with each term multiplied by $\hbar$. This is likely to be confusing when you first encounter it. With $\hbar$ multiplying all the terms, eqn (2.51) takes on the appearance of a momentum conservation equation. However, because $\hbar\boldsymbol{k}$ is not a true momentum, eqn (2.51) is a statement of *wave vector conservation* or *crystal momentum conservation*.

2.12. Phonon-Phonon Scattering

The above considerations set the stage for a discussion of crystal momentum relaxation in a crystal. The idea is this. Unless phonons interact with one another (or with electrons, defects, surfaces, etc.) there can be no dissipation of their crystal momentum, and consequently the thermal conductivity is infinite. In other words, phonons pass freely throughout the crystal. Suppose two phonons, $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$, interact and are annihilated thereby creating a third phonon, $\boldsymbol{k}_3$. If the $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ phonons lie well within the first Brillouin zone, so does the $\boldsymbol{k}_3$ phonon. Figure 13(a) shows how this works. This situation prevails at low temperatures because the phonons have low energy. They occupy mainly the low-k region of the acoustic branches.

Now consider higher temperatures, such that phonons are present all the way to the Brillouin zone edges of the acoustic branches. Suppose two energetic phonons are annihilated to create a third phonon, but the resultant $\boldsymbol{k}_1 + \boldsymbol{k}_2$ lies outside the first Brillouin zone. A phonon cannot be created with this wave vector because phonons do not exist outside the first Brillouin zone. The phonon that is created differs from $\boldsymbol{k}_1 + \boldsymbol{k}_2$ by a reciprocal lattice vector $\boldsymbol{G}$ such that $\boldsymbol{k}_1 + \boldsymbol{k}_2 + \boldsymbol{G} = \boldsymbol{k}_3$, where $\boldsymbol{k}_3$ lies within the first Brillouin zone, as depicted in Fig. 13(b). This is called an umklapp process (German *umklappen*: to turn over) whereas the situation in Fig. 13(a), where the phonons lie within the Brillouin zone, is referred to as a normal process.

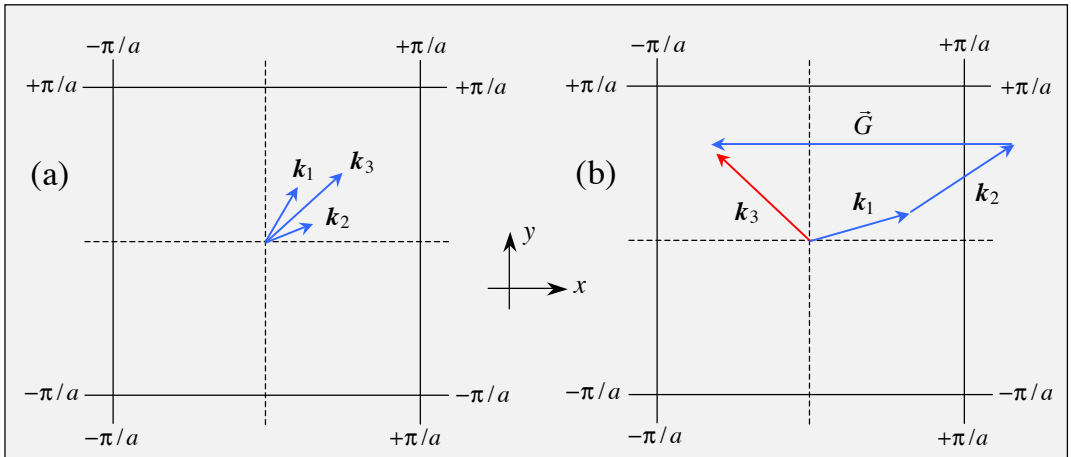


Figure 13. (a) The resultant wave vector lies within the first Brillouin zone. This is a *normal* process. (b) The resultant wave vector lies outside the first Brillouin zone. A reciprocal lattice vector returns the wave vector to the first Brillouin zone. This is an *umklapp* process.

In Fig. 13(b), a simple process is indicated in which $\boldsymbol{G}$ lies along one of the crystal axes. For a 3D crystal, $\boldsymbol{G}$ can be made of components in different directions. For example, suppose $\boldsymbol{k}_1$ and $\boldsymbol{k}_2$ each extend from the origin to near the $+\pi/a, +\pi/a$ corner on the upper right. In this case, a reciprocal lattice vector $\boldsymbol{G} = -(2\pi/a)\hat{x} - (2\pi/a)\hat{y}$ is required to return the system to the first Brillouin zone. Regardless of such details, the important fact is that this umklapp process can dissipate crystal momentum.

2.13. High-Temperature Limit

At high temperatures and with a good quality crystal, umklapp processes dominate crystal momentum relaxation. In this regime, the thermal conductivity $\kappa = Cvl/3$ is proportional to the scattering length l. The heat capacity C is independent of temperature ($C = 3nk_B$), and the speed of sound depends weakly on temperature. The scattering length depends on the phonon density. This is analogous to gas phase kinetics, where the mean distance between collisions varies inversely with number density.

The phonon density depends linearly on temperature, as can be seen by noting the high temperature limit:

$$\langle n_k \rangle = \frac{1}{e^{\omega_k / k_B T} - 1} \quad \Rightarrow \quad k_B T / \omega_k \,. \tag{2.52}$$

Thus, we find that l varies at least as strongly as T^{-1}. The variation can be stronger, depending on multiphonon processes, which are nontrivial to model. In experiments, the thermal conductivity is found to vary as T^{-z}, where $1 \le z \le 2$.

2.14. Low-Temperature Limit

At low temperatures, phonon collisions are dominated by normal processes, which conserve crystal momentum, as indicated in Fig. 13(a). One can say that the umklapp processes are frozen out. This has the interesting consequence that at sufficiently low temperatures there is no mechanism for the thermalization of phonons. In other words, heat can flow freely in the absence of a temperature gradient. Thus, the thermal conductivity of a perfect electrical insulator crystal is infinite in the low temperature limit. Of course we know that this is not the case. As the low temperature limit is approached, the number of phonons that can participate in umklapp processes falls off rapidly with temperature. The phonons that do participate are energetic ($\omega_k \sim k_B \Theta_D$), so the average number of these phonons is

$$\langle n_k \rangle \quad \Rightarrow \quad \frac{1}{e^{\Theta_D / T} - 1} \sim e^{-\Theta_D / T} \tag{2.53}$$

Because the number of effective collisions falls off exponentially as $e^{-\Theta_D / T}$, the mean scattering length (or the mean time between effective phonon collisions) varies as $e^{\Theta_D / T}$ or something close to it, say $e^{T_0 / T}$. As a consequence, defects, impurities, isotopes, and even sample size limit the thermal conductivity. This is illustrated in Fig. 14.

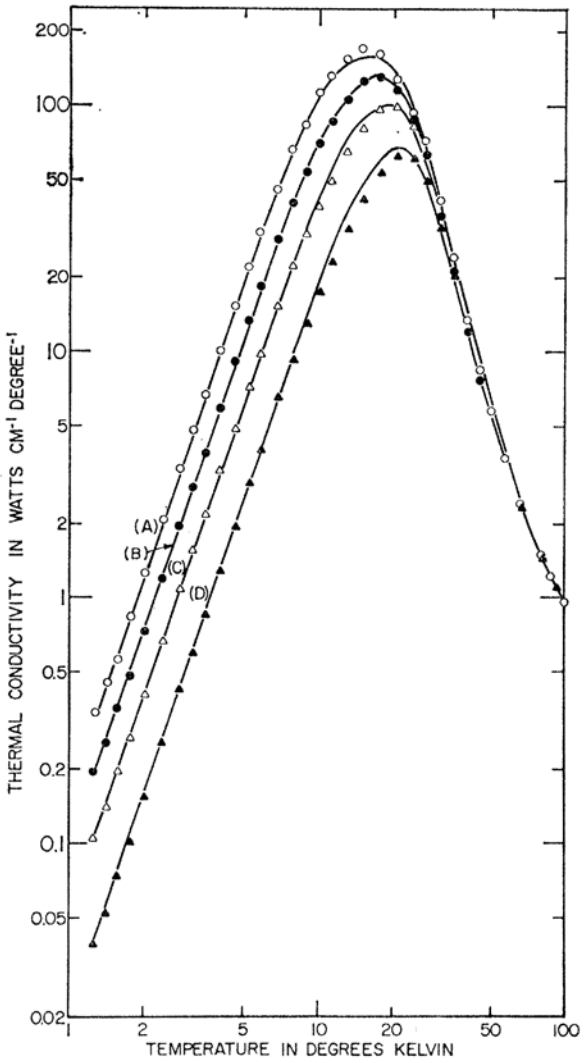


Figure 14. Thermal conductivity κ (in units of $\mathrm{W \cdot cm^{-1} \cdot K^{-1}}$) versus temperature for isotopically pure LiF crystals of different size: Below 10 K, κ varies as T^3. This variation is due to the temperature dependence of the heat capacity in the low temperature regime. Curves (A) – (D) are for crystals having different dimensions (in mm): (A) 7.55 × 6.97; (B) 4.24 × 3.77; (C) 2.17 × 2.10; and (D) 1.23 × 0.91. These crystals are sufficiently free of defects and impurities that these imperfections do not contribute to the thermal conductivity, even at the lowest temperatures. For temperatures below ~ 10 K, the scattering length l is of the order of the size of the crystal. Taken from P. D. Thatcher, Phys. Rev. **156**, 975 (1967).

Identity theft