Lattice vibrations - phonons

So far, we have assumed that the ions are fixed at their equilibrium positions, and we focussed on understanding the motion of the electrons in the static periodic potential created by the ions. But, of course, the ions are quantum objects that cannot be at rest in well-defined positions – this would violate Heisenberg's uncertainty principle. So they must be moving! In the following, we will assume that we are at temperatures very well below the melting temperature of the crystal, so that the ions perform small oscillations about their equilibrium positions. Of course, with increasing T the amplitude of these vibrations increases and eventually becomes large enough that the ions start moving past one another and the lattice order is progressively lost (the crystal starts to melt); our approximations will not be valid once we reach such high temperatures.

Dealing with small oscillations is very nice (as you may remember from studying Lagrangian/Hamiltonian dynamics) because in the harmonic approximation, i.e. where we expand potentials up to second-order in these displacements, such classical problems can be solved exactly by mapping them to a sum of independent harmonic oscillators. Going to the quantum solution is then trivial, because it will simply be the sum of independent quantum harmonic oscillators. In this chapter we discuss this solution and its implications, and for the time being we will completely forget about the electrons. We will return back to them in the next section.

I assume you have already studied phonons in your undergraduate solid state physics course. It is very likely that there you considered some simple cases (chains with 1 atom and 2 atom bases) from which you inferred the general characteristics of the phonons of any solid. Here we will generalize this discussion to any 3D crystal. This makes the discussion a bit more mathematical but it's really the same idea, so try to not be blinded by the math and to see that the underlying physics is very simple, just harmonic motion.

The general case of interest to us is a 3D crystal with a total of $N = N_1N_2N_3$ unit cells and periodic boundary conditions. To be specific, let $\vec{a}_i, i = 1, 3$ be the lattice vectors that define the unit cell. For a simple cubic lattice, they would point along the 3 axes and have equal lengths, but of course much more complicated possibilities exist. In any event, we assume that the crystal's size is $N_1\vec{a}_1 \times N_2\vec{a}_2 \times N_3\vec{a}_3$ with periodic boundary conditions. As always, we will let $N_i \to \infty$ in the end. We index unit cells with $n = (n_1, n_2, n_3)$ where each $n_i = 1, N_i$, and their locations are, therefore, specified by $\vec{R}_n = \sum_{i=1}^{3} n_i\vec{a}_i$.

Let us assume that there are r atoms inside each unit cell, which we will index with $\alpha = 1, r$, and let their masses be M_α and their equilibrium positions (with respect to the reference point of the unit cell) be $\vec{\rho}_\alpha$. These we will assume to be known, in the following. What interests us is the motion about these (known) equilibrium positions.

Then, the location of the α ion of unit cell n is $\vec{R}_{n,\alpha}(t) = \vec{R}_n + \vec{\rho}_\alpha + \vec{u}_{n,\alpha}(t)$, where $u_{n,\alpha}(t)$ is its displacement from its equilibrium position. Until now, we assumed that $\vec{u}_{n,\alpha}(t) \equiv 0$, i.e. ions were frozen at their equilibrium positions. From now on, we'll allow these quantities to be true variables, but their amplitudes are restricted to be much much smaller than the separation from their closest neighbors, so that we still have a well-defined crystal structure.

For simplicity, we will solve this problem first in its classical version, and then quantize it and get the quantum solution. Also, I'll use the lagrangian approach and then switch to the Hamiltonian one, before quantizing. Of course, one can do things in a different order as well and the final answer is the same, but I think this is the simplest way to get to it. Using the $3rN$ coordinates of the Nr ions as the generalized coordinates, the Lagrangian for this corresponding classical problem is:

$$\mathcal{L}_{ions} = \sum_{n=1}^{N} \sum_{\alpha=1}^{r} \frac{M_\alpha \dot{\vec{R}}_{n,\alpha}^2}{2} - \frac{1}{2} \sum_{n,n'} \sum_{\alpha,\alpha'} V_{n\alpha,n'\alpha'} \tag{1}$$

where the second term is the sum of interactions between all pairs of ions. Note that even though I said that we completely ignore the electrons, in fact V is screened by them, in other words we would not use the Coulomb form, but something with a shorter range. Usually, in practice one restricts it to nearest-neighbor pairs only, because of this. The formalism works for any form, though, this is just a matter of how tedious the actual calculations get. Again, we assume this potential to be known.

So far, this is exact. Here we make our **harmonic approximation**, i.e. we switch to the variables $\vec{u}_{n,\alpha}(t)$ and expand the potential to second order in these small displacements. This means that we replace the interactions by a sum of three terms. The first one is the equilibrium contribution, i.e. what we get when we set $\vec{u}_{n,\alpha}(t) = 0$. This is just an overall constant energy which is part of the cohesion energy of the crystal. The first order term is the contribution linear in $\vec{u}_{n,\alpha}(t)$. This must vanish because we are expanding about equilibrium positions, therefore the potential gradients that appear in the coefficients must all vanish. So only the second order correction survives (in this approximation) and we have:

$$\mathcal{L}_{ions} \approx \sum_{n=1}^{N} \sum_{\alpha=1}^{r} \frac{M_\alpha \dot{\vec{u}}_{n,\alpha}^2}{2} - \frac{1}{2} \sum_{n,n'} \sum_{\alpha,\alpha'} \sum_{i,i'} \phi_{n\alpha i}^{n'\alpha' i'} u_{n\alpha i} u_{n'\alpha' i'} \tag{2}$$

where $u_{n\alpha i}, i = 1, 3$ are the cartesian coordinates of the displacement $\vec{u}_{n,\alpha}(t)$ and

$$\phi_{n\alpha i}^{n'\alpha' i'} = \left. \frac{\partial^2 V_{n\alpha,n'\alpha'}}{\partial R_{n\alpha i} \partial R_{n'\alpha' i'}} \right|_{eq}$$

are known quantities, called atomic force constants. In fact, you should think of ϕ as being a $3rN \times 3rN$ matrix that contains all the information regarding the interactions between ions (at this level of approximation). Clearly, it has to satisfy some symmetry conditions. First, all its entries are real numbers, for obvious reasons. Second, it is clearly a symmetric matrix:

$$\phi_{n\alpha i}^{n'\alpha' i'} = \phi_{n'\alpha' i'}^{n\alpha i}$$

To see the origin of the third symmetry requirement, consider the Euler-Lagrange (classical) equations of motion for ion (n, α):

$$M_\alpha \ddot{u}_{n\alpha i} = - \sum_{n',\alpha',i'} \phi_{n\alpha i}^{n'\alpha' i'} u_{n'\alpha' i'}$$

If we displace all atoms equally, i.e. we change all $u_{n'\alpha' i'} \to u_{n'\alpha' i'} + \delta_i$, then the force on ion (n, α) cannot change. From the equation above, this means that:

$$\sum_{n'\alpha'} \phi_{n\alpha i}^{n'\alpha' i'} = 0$$

Similarly, we could rotate the crystal as a whole and that would not change the forces, either. This gives a further set of general constraints (a bit of care is needed here as the cartesian projections of the force do change, as described by a rotation matrix).

Finally, because we do have an underlying lattice structure, we must have:

$$\phi_{n\alpha i}^{n'\alpha' i'} \equiv \phi_{\alpha i}^{\alpha' i'}(|\vec{R}_n - \vec{R}_{n'}|)$$

i.e. because the potential depends on the unit cells to which the ions belong only through the distance between the unit cells, so do its derivatives. Of course, if there are additional symmetries, for example very symmetric locations of like ions inside the unit cell, then this matrix will reflect that as well. In any event, so we know this matrix, and the challenge is to solve the equations of motion written above. We are looking for normal modes (because any general solution can be written as a linear combinations of them); these are solutions where all atoms oscillate with the same frequency. Moreover, because of the lattice structure, we expect solutions to reflect this periodicity (similarly located atoms should exhibit similar motion). So we try solutions of the form:

$$u_{n\alpha i}(t) = \frac{1}{\sqrt{M_\alpha}} e_{\alpha i} e^{i\vec{q}\vec{R}_n - i\omega t}$$

where $e_{\alpha i}$ are the new unknowns. The factor in front is usually pulled out, although of course it's a constant that could be included into the e's. Before continuing, we use the periodic boundary conditions to find the restrictions on the vector $\vec{q}$. Expressing it in terms of the reciprocal lattice vectors $\vec{b}_i, i = 1, 3$ (defined such that $\vec{a}_i\vec{b}_j = 2\pi\delta_{ij}$, this results in allowed momenta $\vec{q} = \sum_i \frac{m_i}{N_i}\vec{b}_i$ where m_i are integers. Since we only have plane waves at discrete locations, we find that only consecutive N_i values of m_i lead to distinct solution, and then they repeat itself. So, in fact, that are N distinct values allowed for $\vec{q}$, lying inside the first Brillouin zone. Hopefully this is not a shock: by now you should expect that anytime we deal with a periodic lattice, all quantities are periodic in (quasi)momentum and therefore its values are restricted to the BZ.

With this, the equations of motion become:

$$\omega^2 e_{\alpha i} = \sum_{\beta j} D_{\alpha i}^{\alpha' i'}(\vec{q}) e_{\alpha' i'}$$

where

$$D_{\alpha i}^{\alpha' i'}(\vec{q}) = \sum_n \frac{1}{\sqrt{M_\alpha M_\beta}} \phi_{\alpha i}^{\alpha' i'}(\vec{R}_n) e^{i\vec{q}\vec{R}_n}$$

are $3r \times 3r$ matrices (known, as well; they are called **the dynamical matrix**), one for each allowed value of $\vec{q}$. So the beauty is that using the lattice's translational invariance, we separated the initial set of $3rN$ coupled equations, into N distinct sets of $3r$ equations each; needless to say, this is huge progress and a nice example why one should always take advantage of symmetries! Now, such homogeneous linear systems have non-trivial solutions only for values ω^2 which are eigenvalues of the matrix D, while the e coefficients are the corresponding eigenvectors, which we choose to be orthonormal. Note that we know that the matrix D is hermitian symmetric because ϕ is symmetric, so we're guaranteed that the eigenvalues are real. It turns out that they are also positive if the structure we chose truly is the equilibrium one (i.e., total energy is minimized, meaning that the matrix of second derivatives is positive defined). So we're guaranteed that ω are real numbers (we always choose the positive solution – this is a convention, it makes no difference which solution is chosen so long as it's only one for each eigenvalue).

To summarize: for each of the N allowed $\vec{q}$, we find $3r$ distinct normal modes with frequencies $\omega_j(\vec{q})$, $j = 1, ..., 3r$. The corresponding eigenvectors $\vec{e}_\alpha^{(j)}(\vec{q})$ then define the spatial profile of the corresponding solution, in other words show in which direction the atoms of type $\alpha = 1, ..., r$ inside each unit cell, are displaced when this normal mode is activated. Because the displacement patterns associated with the normal modes are physical quantities, in other words must be real numbers, it is straightforward to check that we must have $\vec{e}_\alpha^{(j)*}(\vec{q}) = \vec{e}_\alpha^{(j)}(-\vec{q})$. The orthonormation condition also implies that we choose $\sum_\alpha \vec{e}_\alpha^{(j)*}(\vec{q})\vec{e}_\alpha^{(j')}(\vec{q}) = \delta_{j,j'}$.

The general solution, then, is a linear combination of all these normal modes:

$$\vec{u}_{n\alpha}(t) = \sum_{j,\vec{q}} C_j(\vec{q})\vec{e}_\alpha^{(j)}(\vec{q})\frac{e^{i\vec{q}\vec{R}_n - i\omega_j(\vec{q})t}}{\sqrt{M_\alpha N}}$$

where $C_j(\vec{q})$ are $3rN$ complex numbers which define the amplitudes and phases of the various normal modes. Of course, to find these we need to be given the $2 \times 3rN$ initial conditions (initial positions and initial velocities for all atoms). So this would end the classical solution.

Before quantizing, let's rewrite the Lagrangian in terms of new generalized coordinates and speeds, corresponding to these normal modes. In other words, we change variables from the old u's to the new Q's defined such that:

$$\vec{u}_{n\alpha}(t) = \sum_{j,\vec{q}} Q_j(\vec{q},t)\vec{e}_\alpha^{(j)}(\vec{q})\frac{e^{i\vec{q}\vec{R}_n}}{\sqrt{M_\alpha N}}$$

Of course, we already know that the solutions for the new variables will be $Q_j(\vec{q},t) = C_j(\vec{q})e^{-i\omega_j(\vec{q})t}$. The point is that these are so simple, that one can easily guess that the Lagrangian should look a lot simpler written in terms of the Qs. After straightforward calculations using the equations discussed above, one finds indeed that:

$$\mathcal{L}_{ions} = \frac{1}{2}\sum_{j,\vec{q}}\left[\dot{Q}_j(-\vec{q},t)\dot{Q}_j(\vec{q},t) - \omega_j(\vec{q})^2 Q_j(-\vec{q},t)Q_j(\vec{q},t)\right]$$

So yes, we rewrote the messy original Lagrangian as a sum over $3rN$ independent harmonic oscillators. Now let's turn this into a Hamiltonian, and then quantize it. For the first step, we replace the generalized velocities with the generalized momenta $P_j(\vec{q},t) = \frac{\partial \mathcal{L}_{\text{ions}}}{\partial \dot{Q}_j(\vec{q},t)}$ to find:

$$H_{ions} = \frac{1}{2}\sum_{j,\vec{q}}\left[P_j(-\vec{q},t)P_j(\vec{q},t) + \omega_j(\vec{q})^2 Q_j(-\vec{q},t)Q_j(\vec{q},t)\right]$$

Now we quantize by asking the momenta and coordinates to be operators:

$$\hat{H}_{ions} = \frac{1}{2}\sum_{j,\vec{q}}\left[\hat{P}_j(-\vec{q})\hat{P}_j(\vec{q}) + \omega_j(\vec{q})^2 \hat{Q}_j(-\vec{q})\hat{Q}_j(\vec{q})\right]$$

where, as usual $[\hat{Q}_j(\vec{q}), \hat{P}_{j'}(\vec{q'})] = i\hbar\delta_{\vec{q},\vec{q'}}\delta_{jj'}$. Just like we do for ordinary quantum harmonic oscillators, it is convenient to define ladder operators for each mode:

$$\hat{Q}_j(\vec{q}) = \sqrt{\frac{\hbar}{2\omega_j(\vec{q})}}\left(b_{j,\vec{q}} + b^\dagger_{j,-\vec{q}}\right) \qquad \hat{P}_j(\vec{q}) = i\sqrt{\frac{\hbar\omega_j(\vec{q})}{2}}\left(b^\dagger_{j,\vec{q}} - b_{j,-\vec{q}}\right)$$

where $[b_{j,\vec{q}}, b_{j',\vec{q'}}] = 0, [b_{j,\vec{q}}, b^\dagger_{j',\vec{q'}}] = \delta_{\vec{q},\vec{q'}}\delta_{jj'}$ are bosonic operators. In terms of these, we finally find:

$$\hat{H}_{ions} = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q})\left(b^\dagger_{j,\vec{q}}b_{j,\vec{q}} + \frac{1}{2}\right)$$

Here we have the zero point motion energy $E_{GS} = \frac{1}{2}\sum_{j,\vec{q}} \hbar\omega_j(\vec{q})$ which should also be included into the cohesion energy – this contribution is always there, and it is precisely because even in the ground-state, the ions do oscillate and therefore contribute something to the total energy. Excited states are $|\{n_{j,\vec{q}}\}\rangle = \prod_{j,\vec{q}} \frac{[b^\dagger_{j,\vec{q}}]^{n_{j,\vec{q}}}}{\sqrt{n_{j,\vec{q}}!}}|0\rangle$ with energy $E_{ions}(\{n_{j,\vec{q}}\}) = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q})\left(n_{j,\vec{q}} + \frac{1}{2}\right)$. Here, $n_{j,\vec{q}} = 0, 1, 2, ...$ defines the eigenstate of each normal mode (oscillator).

We are now basically done: we know all eigenvalues and eigenfunctions. Of course, in actual calculations we need to go back to the original operators which were the displacements of the ions out of equilibrium. Given the link between us and Qs, and the Qs and the bs, this is easily achieved:

$$\hat{\vec{u}}_{n\alpha} = \sum_{j,\vec{q}} \hat{Q}_j(\vec{q})\vec{e}^{(j)}_\alpha(\vec{q})\frac{e^{i\vec{q}\vec{R}_n}}{\sqrt{M_\alpha N}} = \sum_{j,\vec{q}} \sqrt{\frac{\hbar}{2M_\alpha\omega_j(\vec{q})}}\vec{e}^{(j)}_\alpha(\vec{q})\frac{e^{i\vec{q}\vec{R}_n}}{\sqrt{N}}\left(b_{j,\vec{q}} + b^\dagger_{j,-\vec{q}}\right) \tag{3}$$

and you can similarly find the corresponding momenta. So we can calculate any desired matrix elements of any operators that can be expressed in terms of the ions' locations and momenta.

Let see what **phonons** are. For simplicity, consider the eigenstate $|0, ..., 0, n_{j,\vec{q}} = n, 0, ..., 0\rangle$, i.e. the ground-state for all harmonic oscillators except $j, \vec{q}$ which is in its n^{th} -excited state. Then the energy is $E_{GS} + \hbar\omega_j(\vec{q})n$ because we use the raising operator $b^\dagger_{j,\vec{q}}$ to go up n times from the GS for this oscillator.

Instead of thinking in these terms, we can use the following (totally equivalent, but more convenient) language. We think of the ground-state $|0, 0, ..., 0\rangle = |0\rangle$ as a **vacuum** for excitations, and say that state $|0, ..., 0, n = n_{j,\vec{q}}, 0, ..., 0\rangle$ has n phonons of the mode j and momentum $\vec{q}$, where each such phonon costs an energy $\hbar\omega_{j,\vec{q}}$. So the total energy is again $E_{GS} + \hbar\omega_j(\vec{q})n$ because of these n present phonons. In this language, then, $b_{j,\vec{q}}, b^\dagger_{j,\vec{q}}$ are phonon annihilation and creation operators, because acting with $b^\dagger_{j,\vec{q}}$ on a state will add to it a phonon of type j and momentum $\vec{q}$, etc. So phonons are bosonic particles of $j = 1, ..., 3r$ distinct types, and such that if their momentum is $\vec{q}$ then their energy is $\hbar\omega_{j,\vec{q}}$. If you think about it, this is just like we think about photons – except those are the bosons we get when we quantize the E&M fields and there are only 2 types, corresponding to circularly left and right polarized ones. But just like we say that photons are emitted or absorbed when a system (eg an electron) moves in between two eigenstates, and the photon is such that momentum and energy is conserved; just like that phonons are emitted or absorbed as the lattice switches from one eigenstate to another (because of interactions with something else, of course); and the phonon(s) emitted or absorbed are such that total momentum and energy is conserved. That's it - but it's very convenient to think in this picture.

Each of the $3r$ distinct flavors of phonons is called a **phonon mode**. As already stated, they are characterized by their energies $\hbar\omega_j(\vec{q})$ where $j = 1, ..., 3r$. These energies depend on the momentum $\vec{q}$ of the phonon and are periodic function with BZ periodicity.

Phonon modes are classified in a couple of ways. First, precisely 3 of these modes are so-called **acoustic modes** (correspondingly, we talk about acoustic phonons). By definition, a mode is acoustic if its energy vanishes as $\vec{q} \to 0$; in other words these modes are gapless. In fact, we can be a bit more precise. It turns out that for these modes, $\omega_j(\vec{q}) \to c_j q$ as $q \to 0$, where c_j is the speed of sound for that mode. Of the 3 modes, one is **longitudinal**, i.e. its polarization vectors $\vec{e}^{(j)}_\alpha(\vec{q})$ become parallel to $\vec{q}$ as $q \to 0$ (in other parts of the Brillouin zone the polarizations can point whichever way. We still call the mode "longitudinal"). If you look at the equation for the displacements $\hat{\vec{u}}_{n\alpha}$, this means that in this mode, for small q, all atoms are displaced parallel to $\vec{q}$, hence the name. As you guessed, there are also two **transverse** modes, where the atoms are displaced perpendicular to $\vec{q}$ when $q \to 0$. The speed of sound associated with these three modes is usually different, although in highly symmetric systems the transverse modes will have the same speed of sound. These modes are called acoustic precisely because they are activated when sound propagates through the crystal. These acoustic phonons are also called **Goldstone modes**. There is a **Goldstone theorem** which states that everytime a system breaks a continuos symmetry, there is a gapless boson associated with that. The symmetry broken here is the (continuous) translational symmetry, times 3 as we can translate in the x, y, z directions. The crystal only has discrete translational symmetry, not continuous like free space, so there must be precisely 3 Goldstone modes, and these are the acoustic modes. The way to think about this, is that if we translate all atoms by the same δ_i, $i = x, y, z$, then although we have displaced all atoms, in reality we haven't changed their relative positions. So this should not cost any energy, and this is why these Goldstone modes are gapless. Indeed, as $\vec{q} \to 0$, one can show that the pattern of displacemens $\langle \hat{\vec{u}}_{n\alpha} \rangle$ in a state with only acoustic phonons looks like an "in-phase" motion where all atoms are displaced in the same direction with a wavelike pattern with a wavelength $\lambda \sim 1/q$. As $q \to 0$ this is indeed an overall translation and will cost no energy, hence the gapless modes.

Note that if we have a single atom per unit cell, i.e. $r = 1$, then the 3 resulting modes are the three acoustic modes, and that is all. For more complicated crystals with $r \geq 2$, the remaining $3(r-1)$ modes are called **optical modes**. These are **gapped modes**, and in particular $\omega_j(\vec{q})$ remains finite as $q \to 0$. Again, in this limit we can classify them as longitudinal vs transverse. So we can also say that there are $r-1$ optical modes, each of which has one longitudinal and two tranverse types. These modes are called optical for the following reason. One can show that in the limit $q \to 0$, the various atoms in the unit cell are displaced in anti-phase. To give a simple example, if we have a chain with two atoms per unit cell, then if its longitudinal optical mode is excited, we find that $\langle u_1(t) \rangle \propto -\langle u_2(t) \rangle$. In other words if one (set of) atoms moves right, the other (set) moves left, and viceversa. Remember that these are ions, and they will generally have different ionic charges. So this anti-phase motion gives rise to oscillating dipoles, and those couple very strongly to light (oscillating E&M fields). In other words, these modes can be excited optically, hence their name.

0.1 Phonons - when is this concept truly useful

What we did so far is all exact, once we accept the **harmonic approximation**. So we have to worry what happens if that is not a good approximation and we had to keep higher order terms in the potential's expansion (for instance, because of going to higher temperatures). We can do that by building on what we already have, as we'll see soon, but the point is that since we change the Hamiltonian then we obviously will change the eigenstates, and then we have to wonder whether the notion of phonons survives this change (after all, we introduced them precisely because we had a very specific structure for the eigenenergies, ie the sum of many independent harmonic oscillators).

So let's add cubic corrections to the Hamiltonian - those would be the next biggest contribution. By the way, such higher order corrections are called **anharmonic terms**. Then, the Hamiltonian can be written as:

$$\hat{H}_{ions} = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q}) \left(b^{\dagger}_{j,\vec{q}} b_{j,\vec{q}} + \frac{1}{2} \right) + \frac{1}{6} \sum_{n,n',n''} \sum_{\alpha,\alpha',\alpha''} \sum_{i,i',i''} \phi(...) \hat{u}_{n\alpha i} \hat{u}_{n'\alpha' i'} \hat{u}_{n''\alpha'' i''}$$

where $\phi(...)$ are known coefficients depending on all those indexes, which are obtained as 3rd order partial derivatives of the potential. Of course, next we rewrite the displacement operators in terms of phonon creation and annihilation operators like in Eq. (3) so that we don't mix languages. Then, the Hamiltonian can be written as:

$$\hat{H}_{ions} = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q}) \left(b^\dagger_{j,\vec{q}} b_{j,\vec{q}} + \frac{1}{2} \right) + \frac{1}{6} \sum_{\substack{j_1,j_2,j_3 \\ \vec{q}_1,\vec{q}_2}} M_{j_1,j_2,j_3}(\vec{q}_1,\vec{q}_2,\vec{q}_3) \left(b_{j_1,\vec{q}_1} + b^\dagger_{j_1,-\vec{q}_1} \right) \left(b_{j_2,\vec{q}_2} + b^\dagger_{j_2,-\vec{q}_2} \right) \left(b_{j_3,\vec{q}_3} + b^\dagger_{j_3,-\vec{q}_3} \right)$$

Clearly the three parantheses come from the expansion of the three displacements, and all the details of the interaction are included into the M coefficients. Such terms are called **anharmonic terms**. In particular, they have to encode conservation laws which come from the symmetries of the system. One of this is conservation of the (quasi)momentum because the lattice is invariant to lattice translations. You might assume that this means that $\vec{q}_1 + \vec{q}_2 + \vec{q}_3 = 0$, given that processes described in each paranthesis lower the momentum of a state they act upon by their respective $\vec{q}$, either by removing a phonon with that momentum or by creating one with its opposite momentum. That conservation law would be true in free space, but not for a lattice, i.e. not for a discrete translational group. Remember that because of this discretness, each $\vec{q}$ is contrained to the first BZ so as to not overcount how many distinct allowed momenta we have, given that there is no difference between, say, a phonon with momentum $\vec{q}$ and one with $\vec{q}+\vec{G}$, where $\vec{G}$ is any reciprocal lattice vector. Because of this, if you go through the exercise to calculate the M's, you'll find that the conservation law is $M_{j_1,j_2,j_3}(\vec{q}_1,\vec{q}_2,\vec{q}_3) \sim \sum_{\vec{G}} \delta_{\vec{q}_1+\vec{q}_2+\vec{q}_3,\vec{G}}$, i.e. the total momentum exchanged can be any of the reciprocal lattice vectors – including zero, of course. The way this works in practice is the following. Suppose we keep $\vec{q}_1$ and $\vec{q}_2$ as free variables and wonder what is the allowed $\vec{q}_3$ (with our convention that all these vectors are restricted to the first BZ). Then, if $\vec{q}_1$ and $\vec{q}_2$ are small enough, we can find a value $\vec{q}_3 = -\vec{q}_1 - \vec{q}_2$ that also lies inside the BZ, and all is good, this looks like the usual momentum conservation. However, if $\vec{q}_1$ and $\vec{q}_2$ are large and pointing sort of in the same direction, then $-\vec{q}_1 - \vec{q}_2$ will lie outside the BZ. In this case, there is precisely one possible $\vec{G}$ such that $\vec{q}_3 = \vec{G} - \vec{q}_1 - \vec{q}_2$ is returned back to the first BZ – and this is the $\vec{G}$ that appears in the (quasi)momentum conservation. By the way, processes where $\vec{G} \neq 0$ are called **Umklapp processes**.

So let's now consider what sort of processes are described by this additional term. We have three-phonon creation or annihilation processes, which go like $b^\dagger_1 b^\dagger_2 b^\dagger_3$ or $b_1 b_2 b_3$. Their contributions turn out to be small because of energy conservation – basically that makes them quite unlikely to happen (but they are not forbidden!). The more likely processes, which contribute more substantially, are of the type $b^\dagger_1 b^\dagger_2 b_3$ or $b^\dagger_1 b_2 b_3$, i.e. either one phonon is absorbed and two are created in its place, or viceversa (you can check that quasimomentum is always conserved). One thing we deduce from this is that the number of phonons is not conserved, so in this sense they are not like **real particles**, eg electrons, which cannot appear or disappear out of nowhere. (Of course, this statement is not quite true. There are vacuum fluctuations that can create electron-positron pairs, but then we have the positron to deal with and include in our models, but more importantly these are very high energy processes which we are irrelevant at the low-energy scale of interest to us). Because of this non-conservation, **the chemical potential for phonons is zero**. One way to think about it is that we introduce a chemical potential when we want to switch from a canonical description (with a fixed number of particles) to a grand-canonical one (with a variable number) $\mathcal{H} \to \mathcal{H} - \mu\hat{N}$. If the number is variable no matter what, then there is no difference between the two formulations, and for that to be true we must have $\mu = 0$. We'll see below the consequences of this.

The more worrisome conclusion is that we cannot actually think of the state of our lattice in terms of it having so many phonons of such and such types, if the number of phonons and their types keep changing. This should remind you of the discussion we had about quasiparticles, about when it is meaningful to think in terms of them occupying a Fermi sea. It is precisely the same here: if the anharmonic terms are so small that this scattering and changing of the phonons is very very slow, then we can ignore it and it is perfectly useful and fine to think in terms of phonons. If the anharmornic terms are large and these processes happen on a time-scale similar to whatever process is of interest to us, then a description in terms of phonons is meaningless and we have to do better. The good news is that for all reasonable materials at reasonable temperatures (i.e, well below the melting point) the phonon lifetimes are long enough that it is perfectly ok to think in terms of them as excitations with (essentially) infinite lifetimes.

0.2 The lattice contribution to various properties of the solid

Now we can discuss the contribution of the ions to properties of the solid. I will only do so briefly because this should have been discussed in solid state/stat mech courses, so I'll just remind you how this works. Like for any bosons with $\mu = 0$, in thermal equilibrium the average number of phonons of type j and momentum $\vec{q}$ is given by the Bose-Eistein distribution:

$$\langle b^{\dagger}_{j,\vec{q}} b_{j,\vec{q}} \rangle = \frac{1}{e^{\beta \hbar \omega_j(\vec{q})} - 1}$$

This tells us that at "low temperatures", which here means $k_b T \ll \hbar\omega_j(\vec{q})$, there are basically no phonons of that type, whereas at high temperatures their number increases linearly with temperature, $\langle b^{\dagger}_{j,\vec{q}} b_{j,\vec{q}} \rangle \to \frac{k_B T}{\hbar \omega_j(\vec{q})}$. Note that what is "high" and what is "low" temperature now depends very much on the type of phonons: oftentime temperatures of interest are "high" for acoustic phonons but "low" for optical phonons, as the latter have much higher energies than the former. You might also worry what happens with the $\vec{q} = 0$ acoustic phonons, of which it seems we have an infinite number no matter what. It turns out they always fall out of calculations, as I'll show in you in the next example. So we could just remove them from consideration, if that makes you feel better. Physically, this is because they describe actual lattice translations, as we discussed, and those cannot contribute to any physical property.

The lattice contribution to the energy of the system is, then:

$$E_{ions} = \sum_{j,\vec{q}} \hbar\omega_j(\vec{q}) \left[\frac{1}{e^{\beta \hbar \omega_j(\vec{q})} - 1} + \frac{1}{2} \right]$$

Once we know the phonon energies, this can be calculated. Note that at temperatures that are high for all phonon modes, we find this energy to be $3rNk_BT$. This is **the Dulong-Petit law** which follows from the equipartition theorem for a classical system. Of course, at high enough T a classical description of our quantum system must become valid.

In fact, the measurable quantity is the specific heat, i.e. how much energy we must give to the system to change its temperature, eg at constant volume. The contribution from the lattice (remember that there is a separate one, due to the electrons) is:

$$C_V = \frac{\partial E_{ions}}{\partial T}$$

This is a rather messy expression, but it can be calculated. Of course, in practice we take $N \to \infty$, i.e. the sums over allowed momenta are replaced by integrals. One can also make various approximations (like Debye or Einstein approximations) for the phonon dispersions $\hbar\omega_j(\vec{q})$ to further simplify things so that those integrals can be evaluated easily. You should have seen all this discussed before. If not, please look at any standard textbook – it's basically just a matter of doing some integrals.

At this point we could stop and move on, but I'd like to show you another way to calculate phonon spectra, where we do the calculation at a quantum level all the way. Of course, it gives precisely the same answers as the classical calculation that was then quantized, that we discussed here. This will give us a chance to practice equation-of-motion techniques for diagonalizing bosonic Hamiltonians, and an explicit example of acoustic and optical modes, which should better illustrate the general stuff I talked about so far. We will chose an example that is simple enough that we can do it fast, but general enough that you should be able to generalize it to basically any other case of interest.

0.3 Calculating the longitudinal phonons of a chain with a two-site basis

Let's consider a chain of alternating ions with masses m and M, respectively. Let the lattice constant be a. Then we can choose the equilibrum position of ions of mass m to be at $x_n = na$, $n = 1, N$ where N is the number of unit cells (and surprise surprise, we'll use PBC again). The ions of mass M are located $X_n = na + \delta$, where δ is a known distance. In the simplest case we would have $\delta = a/2$ so that each ion is halfway between two ions of the other type, but in general that doesn't have to be true.

Let u_n, U_n be the displacements out of equilibrium of these ions. We'll only consider displacements along the chain, so only the longitudinal modes are being discussed. If we make the harmonic approximation and if we restrict the interactions to only those of nearest-neighbor pairs, the resulting Hamiltonian is:

$$\hat{H} = \sum_n \left[\frac{P_n^2}{2M} + \frac{p_n^2}{2m}\right] + \sum_n \left[\frac{k_1(U_n - u_n)^2}{2} + \frac{k_2(U_n - u_{n+1})^2}{2}\right]$$

I should put hats on all the P, p, u, U because they're operators, but it's too bothersome, so bear with me. In the second term, I allowed for different interactions between ions in the same unit cells, which are δ apart, and ions from neighboring cells, which are $a - \delta$ apart. All the details are inside the elastic constants $k_1, k_2 > 0$. Of course, they must be positive because we expanded about an equilibrium state.

At this point we could go back to the classical equations of motion and follow the general recipe we discussed and solve the problem - I bet you've done this in a previous course, but maybe you should practice just to make sure you can do it! After all, it's just solving a classical problem and it shouldn't give you any trouble.

As I said, we'll do this quantum mechanically all the way. First step is to rewrite $\hat{H} = \hat{H}_1 + \hat{H}_2 + \hat{H}_{12}$, where:

$$\hat{H}_1 = \sum_n \left[\frac{P_n^2}{2M} + \frac{(k_1 + k_2)U_n^2}{2}\right]$$

only involve type 1 of ions,

$$\hat{H}_2 = \sum_n \left[\frac{p_n^2}{2M} + \frac{(k_1 + k_2)u_n^2}{2}\right]$$

only involve type 2, and

$$\hat{H}_{12} = -\sum_n \left[k_1 u_n U_n + k_2 U_n u_{n+1}\right]$$

involve both. If we didn't have the last part, the problem would be trivial because we are summing over independent Harmonic oscillators. Let $\Omega^2 = (k_1 + k_2)/M$ and $\omega^2 = (k_1 + k_2)/m$ be the frequency of these oscillations for the two types of ions. Defining the usual raising and lowering operators:

$$\hat{U}_n = \sqrt{\frac{\hbar}{2M\Omega}}\left(B_n + B_n^\dagger\right) \qquad \hat{P}_n = i\sqrt{\frac{\hbar M\Omega}{2}}\left(B_n^\dagger - B_n\right)$$

and

$$\hat{u}_n = \sqrt{\frac{\hbar}{2m\omega}}\left(b_n + b_n^\dagger\right) \qquad \hat{p}_n = i\sqrt{\frac{\hbar m\omega}{2}}\left(b_n^\dagger - b_n\right)$$

we find right away that:

$$\hat{H}_1 = \sum_n \hbar\Omega\left[B_n^\dagger B_n + \frac{1}{2}\right] = \sum_q \hbar\Omega\left[B_q^\dagger B_q + \frac{1}{2}\right]$$

and

$$\hat{H}_2 = \sum_n \hbar\omega\left[b_n^\dagger b_n + \frac{1}{2}\right] = \sum_q \hbar\omega\left[b_q^\dagger b_q + \frac{1}{2}\right]$$

where for the second set of sums we Fourier transformed to momentum space:

$$b_q^\dagger = \sum_n \frac{e^{iq(na)}}{\sqrt{N}} b_n^\dagger \qquad B_q^\dagger = \sum_n \frac{e^{iq(na+\delta)}}{\sqrt{N}} B_n^\dagger$$

Note that I used the actual equilibrium positions in the Fourier exponentials. I could remove the δ from the second definition – that is just an overall change in the phase and should have no effect on the final results (you might want to check that statement). Of course, the PBC restrict the allowed q to be of the form $2m\pi/N$, where $m = -N/2+1, ..., N/2$ places them inside the first BZ.

Note that if we could ignore $\hat{H}_{12}$, then we would be done: both these terms are diagonal, and the answer would be two optical phonons with frequencies ω, Ω, i.e. two Einstein modes. The reason the acoustic mode is missing is because if we ignore $\hat{H}_{12}$ then, in reality, we ignore the interactions between atoms, as here each one is sitting in its own harmonic well. No surprise that sound could not propagate in such a system, because that requires atoms to push each other around to create the propagating wave. But in any event, this gives you an idea of how Einstein models can arise (if you don't remember what those are, I'll remind you at the end of this section).

So let's see what's the effect of the interaction terms between ions. First, we have to do a bit of work to write $\hat{H}_{12}$ in terms of these operators. The end result is:

$$\hat{H} = \sum_q \left[\hbar\omega b_q^\dagger b_q + \hbar\Omega B_q^\dagger B_q - f(q)(b_q + b_{-q}^\dagger)(B_{-q} + B_q^\dagger)\right]$$

where $f(q) = \frac{\hbar}{2\sqrt{mM\omega\Omega}}(k_1 e^{-iq\delta} + k_2 e^{iq(a-\delta)})$. By the way, why are there no Umklapp processes here? Also, I will ignore the constant terms which are absorbed into the cohesion energy.

Now, remember the equation of motion idea. We're aiming to put this in its diagonal form:

$$\hat{H} = \sum_q \sum_{j=1}^{2} \hbar\omega_j(q) b_{j,q}^\dagger b_{j,q}$$

The question is what are the new energies and the new operators $b_{1,q}, b_{2,q}$ and how are they related to the old b_q, B_q. For the diagonal form:

$$\hbar\omega_j(q) b_{j,q} = [b_{j,q}, \hat{H}]$$

(the new operators are also bosonic). Now, the most general link between the new and old operators is:

$$b_{j,q} = \alpha b_q + \beta B_q + \gamma b^\dagger_{-q} + \delta B^\dagger_{-q}$$

where, of course, the coefficients will depend on j, q but I won't write that. You might think that the new b could depend on all the old $b, b^\dagger, B, B^\dagger$, not just these four. But momentum conservation comes to our rescue: only these 4 guarantee that a momentum q is removed from the system when $b_{j,q}$ acts on it. So the equation above (called the equation of motion) becomes, in terms of the old operators:

$$\hbar\omega_j(q)\left(\alpha b_q + \beta B_q + \gamma b^\dagger_{-q} + \delta B^\dagger_{-q}\right) = [\alpha b_q + \beta B_q + \gamma b^\dagger_{-q} + \delta B^\dagger_{-q}, \hat{H}]$$

and the problem is reduced to finding the coefficients and energies which satisfy this. We could proceed with this, or we can further simplify as follows. For instance, if I want to pick only the α term on the lfs, I should take another commutator with $b^\dagger_q$. In other words:

$$\hbar\omega_j(q)\alpha = [[\alpha b_q+\beta B_q+\gamma b^\dagger_{-q}+\delta B^\dagger_{-q}, \hat{H}], b^\dagger_q] = \alpha[b_q, \hat{H}], b^\dagger_q]+\beta[B_q, \hat{H}], b^\dagger_q]+\gamma[b^\dagger_{-q}, \hat{H}], b^\dagger_q]+\delta[B^\dagger_{-q}, \hat{H}], b^\dagger_q]$$

The double commutators are simple-enough to calculate numbers, so this is the first of 4 linear equations that will give us an eigenvalue problem. I hope this reminds you a bit of what we did in the HFA (except there it was an approximation; this is exact because the Hamiltonian is quadratic).

Here is what I get after doing the commutators:

$$\begin{aligned}
\hbar\omega_j(q)\alpha &= \quad \hbar\omega\alpha - f(q)(\beta-\delta) \\
\hbar\omega_j(q)\beta &= \quad \hbar\Omega\beta - f(-q)(\alpha-\gamma) \\
\hbar\omega_j(q)\gamma &= \quad -\hbar\omega\gamma - f(q)(\beta-\delta) \\
\hbar\omega_j(q)\delta &= \quad -\hbar\Omega\delta - f(-q)(\alpha-\gamma)
\end{aligned}$$

Clearly, the desired $\omega_j(q)$ are the eigenvalues, and $(\alpha, \beta, \gamma, \delta)$ is the corresponding eigenvector, for the matrix on the rhs. This should puzzle us, because it seems to suggest 4 different solutions, not 2 like we expect. But remember that even in the other formulation we got twice as many solutions as expected, but we discarded the negative ω values. We have exactly the same deal here: we keep only the positive eigenvalues.

Adding and subtracting the first and third equation, we find $\hbar\omega_j(q)(\alpha-\gamma) = \hbar\omega(\alpha+\gamma)$ and $\hbar\omega_j(q)(\alpha+\gamma) = \hbar\omega(\alpha-\gamma) - 2f(q)(\beta-\gamma)$, which combine into:

$$\left[(\hbar\omega_j(q))^2 - (\hbar\omega)^2\right](\alpha-\gamma) + 2f(q)\hbar\omega(\beta-\gamma) = 0.$$

Same tricks for the other 2 equations lead to:

$$2f(-q)\hbar\Omega(\alpha-\gamma) + \left[(\hbar\omega_j(q))^2 - (\hbar\Omega)^2\right](\beta-\gamma) = 0.$$

The phonon energies are now easily recovered from the resulting quadratic equation. Discarding the negative values, we find:

$$\omega_{\pm}(q) = \sqrt{\frac{1}{2}\left[\omega^2 + \Omega^2 \pm \sqrt{(\omega^2+\Omega^2)^2 - 16\omega_0^4 \sin^2\frac{qa}{2}}\right]}$$

where $\omega_0^4 = \frac{k_1 k_2}{mM}$ is also a frequency.

Let's look at the $-$ solution first. For $q \to 0$ we find $\omega_-(q) \to \frac{\omega_0^2 a}{\sqrt{\omega^2+\Omega^2}} q$, so this is the acoustic mode. It increases monotonicaly and reaches its maximum value at the Brillouin zone edges $q = \pm\frac{\pi}{a}$. If you calculate the eigenvectors you can check that the ions move in phase, for $q \to 0$.

The other mode is the optical mode, since $\omega_{\pm}(0) = \sqrt{\omega^2 + \Omega^2}$ is finite. This energy decreases with increasing q, but it is easy to verify that it is above the acoustic mode for all energies. So the phonon spectrum is like that sketched on the left, in the next figure.

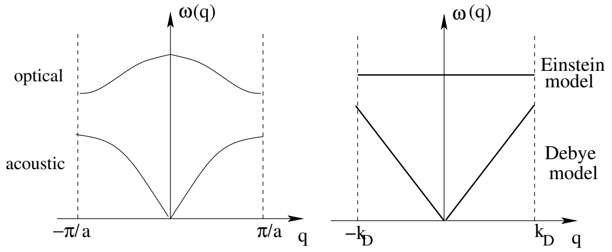

Figure 1: Left: energy of the acoustic and optical longitudinal phonons, $\omega_{\pm}(q)$, for the chain with two atoms in the unit cell. Right: simplified Debye model for the acoustic phonons, assuming linear dispersion at all q but in a spherical Brillouin zone; and Einstein model for the optical phonons, assuming the same energy for all q.

At this point, you should be curious to verify that if you set $m = M, k_1 = k_2$, i.e. make all ions identical and equally spaced, you recover the spectrum for a chain with a one-ion basis, which should only have the acoustic mode! Does that work out?

On the right of the figure, I have sketched for you the ideas behind the Debye and the Einstein approximation. The latter is straightforward, we simply take the optical phonon frequency to be independent of q. That simplies calculations tremendously, of course. It works well because usually those bands are fairly narrow, i.e. nearly constant.

The Debye approximation is for acoustic modes, and the idea is to replace the true dispersion with its long wavelength expression $c_s q$, where c_s is the speed of sound for that mode. Again, this simplifies calculations quite a bit. The logic behind this is that if the temperature is low enough that only phonons with small q are excited in large numbers, then we only have to worry about describing their energies well – and this is what the approximation does. If we're at temperatures well above the highest value, then this mode basically behaves like a classical harmonic oscillator and we get the right results (equipartition theorem and all that) again. If T is comparable to the phonon energies at the top of the band, then we better use the correct expression if we want to be accurate.

In 3D, the Debye approximation has a second component. Remember that there the BZ has some non-spherical shape, eg it can be a cube. This leads to a problem, because $\hbar\omega(\vec{q}) = c_s q$ is only useful if we can integrate in spherical coordinates, but we can't do that in a cubic (let's say) domain. So then one deforms the BZ to have the same volume but a spherical shape, so that the integral can be done easily (at low-T the phonons near the BZ edges are not excited so they don't contribute to quantities of interest, so it doesn't matter how accurately or badly we model them). We need to keep the same volume because remember that the BZ volume is such that precisely N distinct $\vec{q}$ are allowed, and they have a density that is set by the lattice constant and system size. To get the correct number of solutions thus requires us to keep the same overall volume. Doing this allows one to define a **Debye vector** q_D, which is the radius of this "spherical BZ". For instance, if the true Brillouin zone is cubic and of size $\frac{2\pi}{a} \times \frac{2\pi}{a} \times \frac{2\pi}{a}$, then we must have $\frac{4\pi q_D^3}{3} = \left(\frac{2\pi}{a}\right)^3$. One then also talks about a Debye frequency, which is $\hbar\omega_D = c_s q_D$, i.e. the maximum value allowed for these phonons in this approximation. This is a good approximation if $k_B T \ll \hbar\omega_D$ or if $k_B T \gg \hbar\omega_D$.

Now we can go back and discuss the influence of this lattice motion on the electronic behavior.