Second quantization

Gabriel T. Landi
University of São Paulo
November 8, 2019

Contents

1 Introduction

Second quantization is the name given to a set of techniques which are incredibly powerful for dealing with **systems composed by a large number of particles**. It is, in a sense, a different posture towards quantum mechanics. It is still quantum mechanics, but looked at in a different way. The usual formulation of quantum mechanics in terms of the *tensor product* amounts to specifying in which state each particle is. If we have 10^{23} particles this becomes too cumbersome. Moreover, if the particles are all indistinguishable, this doesn't even make sense anymore because we cannot say *which* particle is in which state. All we can say is **how many** particles are in each state. This is called **occupation number representation** and is the basic idea behind second quantization.

Second quantization is extremely powerful. We will learn in these notes to think about second quantization as a *language*. It is a different way of seeing nature. And this will provide us with a very nice way of unifying different fields of knowledge. Condensed matter, quantum optics, statistical mechanics and high energy physics, can all be cast in a similar language using second quantization. By learning this language, you will therefore be able to better navigate between these different fields.

Introducing second quantization for the first time can also be quite confusing. To learn it, gaining intuition is essential. If I try to teach you rigorously from the start, you will be very confused. Here I will therefore focus on an informal introduction to the subject, which will allow us to jump right into applications and therefore gain intuition. If you are curious on how to put this on more rigorous grounds, I recommend Feynman's *Statistical Mechanics*, chapter 7. The downside of doing an informal introduction is that you may be left with the feeling that second quantization is somewhat mystical. It is not. Second quantization *is* quantum mechanics. It is just written in a different way. Please remember that.

2 Single-particle states

The basic idea behind the theory is the notion of single-particle states. That is, the set of quantum states that can be occupied by a single particle. This is what will then be used as the building block for states with multiple particles.

2.1 Free particles

Consider a free particle in one dimension, with Hamiltonian

$$H = \frac{p^2}{2m} \tag{2.1}$$

Here and henceforth we set $\hbar = 1$. To diagonalize H it suffices to diagonalize p. We label the eigenvalues and eigenvectors as

$$p|k\rangle = k|k\rangle. \tag{2.2}$$

Our goal is to find the eigenvalues k and corresponding eigenvectors $|k\rangle$. Working in the coordinate representation, in terms of the basis $|x\rangle$, we define wavefunctions $\phi_k(x)$

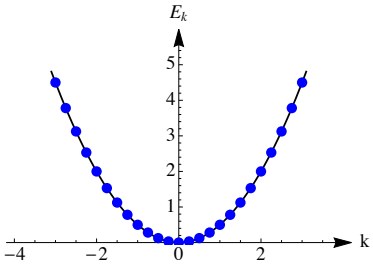

Figure 1: Dispersion relation (2.10). The values of k are actually discrete [Eq. (2.9)] but the discreteness is proportional to $1/L$ and thus becomes very fine when L is the large. The levels therefore form a quasi-continuum

as[1]

$$\phi_k(x) = \langle x|k\rangle, \tag{2.4}$$

so that $|k\rangle$ is expanded as

$$|k\rangle = \int \mathrm{d}x\, |x\rangle\langle x|k\rangle = \int \mathrm{d}x\, \phi_k(x)|x\rangle. \tag{2.5}$$

Since $[x, p] = i$, one may show (as done in any quantum mechanics book) that the eigenstuff equation (2.2) in coordinate representation becomes the differential equation

$$-i\partial_x\phi_k = k\phi_k, \tag{2.6}$$

whose solutions are plane waves $\phi_k = e^{ikx}$.

It is extremely convenient to assume that the system is enclosed in a box of length L and subject to **Periodic Boundary Conditions (PBC)**. This means that $\phi_k(x + L) = \phi_k(x)$. The eigenstates are then written as

$$\phi_k(x) = \frac{e^{ikx}}{\sqrt{L}}, \tag{2.7}$$

where the constant is introduced so that the functions are properly normalized

$$\int_0^L \mathrm{d}x\, |\phi_k(x)|^2 = 1. \tag{2.8}$$

The allowed values of k are determined by imposing the PBC condition $\phi_k(x + L) = \phi_k(x)$. in Eq. (2.7). This leads to $e^{ikL} = 1$. Hence, k is allowed to take on any value $k = \frac{2\pi\ell}{L}$, where $\ell = 0, \pm1, \pm2, \ldots$.

[1] The position kets $|x\rangle$ are a bit different because they are continuous. Orthogonality and completeness must therefore be replaced with

$$\langle x|x'\rangle = \delta(x - x'), \qquad \int \mathrm{d}x\, |x\rangle\langle x| = 1. \tag{2.3}$$

But other than that, they function pretty much like normal kets.

To summarize, the eigenvectors of the momentum operator in a box of size L with PBCs are plane waves of the form (2.7), whereas the eigenvalues are

$$k = \frac{2\pi\ell}{L}, \qquad \ell = 0, \pm 1, \pm 2, \ldots. \tag{2.9}$$

If you want to put $\hbar$ back, you can write the eigenvalues as $\hbar k$ instead. The eigenvalues of the free particle Hamiltonian (2.1) are then simply

$$E_k = \frac{k^2}{2m}. \tag{2.10}$$

Eq. (2.10) is called a **dispersion relation**, or **energy-momentum relation**. It says how energy scales with momentum. The eigenvalues k in Eq. (2.9) are discrete; but their spacing is proportional to $1/L$ and thus becomes very fine when L is large. As a consequence, the spectrum forms a quasi-continuum. The dispersion relation (2.10) is shown in Fig. 1.

$\hbar = 1$ and other dispersion relations

Energy is related to frequency and momentum is related to wavenumber:

$$E = \hbar\omega, \qquad p = \hbar k. \tag{2.11}$$

Thus, when $\hbar = 1$, energy = frequency and momentum = wavenumber, so feel free to use these words interchangeably. This is all that $\hbar$ does. For instance, in electromagnetic theory we find that frequency and wavenumber are related by $\omega = c|k|$, where c is the speed of light. Multiplying by $\hbar$ on both sides yields

$$E = c|p|. \tag{2.12}$$

This is also a dispersion relation, although it is dramatically different from (2.10). We say (2.10) is a **non-relativistic dispersion relation** whereas (2.12) is called a **massless dispersion**.

In relativity we also find things which are somewhere in the middle:

$$E = \sqrt{m^2c^4 + p^2c^2} \tag{2.13}$$

When $pc \ll mc^2$ this yields (2.10) (up to a constant) and when $pc \gg mc^2$ this yields (2.12). In solids we also find much more exotic dispersion relations, which stem due to the confinement of electrons in the potentials created by the atoms. We will see one example below in Sec. 2.2

Generalization to d dimensions

The eigenstates $|k\rangle$ (or the corresponding eigenfunctions ϕ_k) are what we call a **single-particle state**. They are states which completely characterize a particle in 1D. A particle in 3D can be labeled by single-particle states of the form

$$|\boldsymbol{k}\rangle = |k_x, k_y, k_z\rangle = |k_x\rangle \otimes |k_y\rangle \otimes |k_z\rangle, \tag{2.14}$$

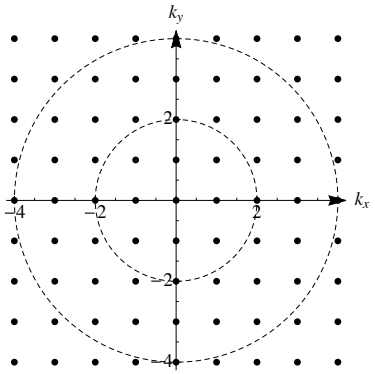

Figure 2: Allowed values for the momentum (k_x, k_y) in 2D [c.f. Eq. (2.9)]. The energy (2.15) correspond to concentric circles in k space.

which are eigenstates of p_x, p_y and p_z. Since momenta in different directions commute, each operator can be diagonalized independently. Thus, each k_i will be discrete exactly like in Eq. (2.9). In 2D this is illustrated in Fig. 2, which shows a grid of (k_x, k_y) values. The dispersion relation (2.10) in d dimensions simply becomes

$$E_k = \frac{\boldsymbol{k}^2}{2m} = \frac{k_1^2 + \ldots + k_d^2}{2m}. \tag{2.15}$$

In 2D, for instance, this corresponds to concentric circles in k space (Fig. 2).

The particle may also have **internal degrees of freedom**. For instance, the particle itself may be an atom, which has internal electronic energy levels. Or, even more simply, the particle may just be an electron which also carries spin. A set of single-particle states for a spin 1/2 free particle is therefore

$$|\boldsymbol{k}\rangle = |k_x\rangle \otimes |k_y\rangle \otimes |k_z\rangle \otimes |\sigma\rangle, \tag{2.16}$$

where $\sigma = \pm 1$.

The important message that you should remember is that *a single particle state completely characterizes the state of a single particle*. I know it sounds redundant. But please remember it! The set of single particle states do not have to be eigenstates of anything. For instance, if we introduce a potential $V(x)$ in the Hamiltonian (2.1), the ϕ_k will no longer be eigenstates. Notwithstanding, they continue to be single-particle states. In fact, a much simpler set of single-particle states are the position eigenstates $|x\rangle$. They also form a basis and they also completely characterize a particle in 1D.

2.2 Tight-binding model

Tight-binding is a simplified toy model which is absolutely lovely. Electrons in semiconductors and metals tend to be trapped near the atomic cores. They can, however, move around by tunneling to neighboring atoms. The motion is therefore a sequential composition of tunnelings. This is what we call **hopping** (Fig. 3). The tight-binding provides a simplified picture of this, where the particles can only like on a **discrete set of states** $|n\rangle$, with $n = 1, 2, \ldots, N$, which are therefore the single-particle states. If the lattice sites are spaced by a fixed amount a (the *lattice spacing*), then each site n is associated with a position $x_n = an$.

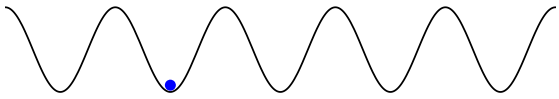

Figure 3: Basic idea behind the tight-binding model, showing a particle hopping through a lattice.

The effects of hopping can be introducing by considering a Hamiltonian

$$H = -g \sum_{i=1}^{N} \Big\{ |n\rangle\langle n+1| + |n+1\rangle\langle n| \Big\}, \tag{2.17}$$

where g is a constant. This Hamiltonian also assumes PBC since the sum goes up to $n = N$. This means that there is a term $|N\rangle\langle N+1|$ which should be interpreted as $|N\rangle\langle 1|$. When a term like $|n+1\rangle\langle n|$ acts on a state $|n\rangle$ it produces $|n+1\rangle$. Thus, $|n+1\rangle\langle n|$ describes a jump from n to $n+1$ (you read it from right to left). Of course, since the dynamics is unitary, one must also be able to jump from $n+1$ to n with the same amplitude.

The Hilbert space in the case of the tight-binding model is finite and has size N. We may thus write Eq. (2.17) as a big matrix with entries $H_{nm} = \langle n|H|m\rangle$. For instance, in the case $N = 5$, it will look like

$$H = -g \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.18}$$

Notice the lonely terms in the upper-right and lower-left corners. These are actually the PBS. In general I do not recommend writing H as a big matrix. We are not interest in $N = 5$, but rather on arbitrary N, so that the matrices would get very big.

We are going to find the eigenvalues and eigenvectors of the tight-binding Hamiltonian. (2.17) for any size N. But before doing so, it is useful to notice how the same idea can be readily extended to more general geometries. A more general tight-binding Hamiltonian will always have the structure

$$H = -\sum_{n,m} g_{nm} |n\rangle\langle m|, \tag{2.19}$$

for some coefficients g_{nm} satisfying $g_{nm} = g^*_{mn}$. We can use this, for instance, to set up a complicated hopping problem through a complicated network of sites, each hopping having a different magnitude and so on and so forth. In the special case where the lattice has **translation invariance**, then $g_{nm} = g(|x_n - x_m|)$ is a function only on the distance between two sites. This is the case of the simple 1D lattice we are using in Eq. (2.17).

Let us now diagonalize the Hamiltonian (2.17). We are looking for the solution of

$$H|\phi\rangle = E|\phi\rangle, \tag{2.20}$$

where the eigenstates $|\phi\rangle$ can be expressed in the basis $|n\rangle$ as $|\phi\rangle = \sum_n \phi_n |n\rangle$. To have a look on what happens when we plug this in Eq. (2.20), let us use the matrix represen-

tation in Eq. (2.18). Then

$$H|\phi\rangle = -g\begin{pmatrix}\phi_2+\phi_N\\ \phi_1+\phi_3\\ \phi_2+\phi_4\\ \phi_3+\phi_5\\ \phi_4+\phi_1\end{pmatrix} = E\begin{pmatrix}\phi_1\\ \phi_2\\ \phi_3\\ \phi_4\\ \phi_5\end{pmatrix} = E|\phi\rangle.$$

We clearly see a pattern here. Each line of the above formula gives us an *algebraic* relation for the coefficients ϕ_i. Moreover, we can clearly guess what is the general form of this equation:

$$\boxed{-g(\phi_{n-1}+\phi_{n+1}) = E\phi_n.} \tag{2.21}$$

Notice how this also works for the boundary terms, provided we use the PBC recipe $\phi_{N+1} = \phi_1$. This therefore converted an eigenvalue/eigenvector problem into a recurrence relation problem. Notice that the variables here are both the vectors ϕ_n as well as the energies E.

The recurrence Eq. (2.21) can be solved by introducing plane waves

$$\phi_n^k = \frac{e^{ikx_n}}{\sqrt{N}}, \tag{2.22}$$

where $x_n = an$ and k is just a label for the different plane waves. The allowed values of k will be determined in a second. I often set the lattice spacing to be $a = 1$, in which case k becomes dimensionless. But if we keep a, then k will have units of 1/length; i.e., wavenumber.

The normalization constant is introduced so that

$$\sum_{n=1}^{N} |\phi_n^k|^2 = 1, \qquad \forall k. \tag{2.23}$$

Plugging Eq. (2.22) into Eq. (2.21) we find that

$$-g\frac{e^{ikn}}{\sqrt{N}}(e^{ik}+e^{-ik}) = E\frac{e^{ikn}}{\sqrt{N}}.$$

Cancelling out the exponentials, we find that the ansatz (2.22) indeed works, but only provided that the energy E_k associated to the plane wave $|\phi_k\rangle$ are given by

$$E_k = -2g\cos ka.$$

This will be the dispersion relation for this model. We will analyze it in more detail in a second.

But before doing so, let us determine what are the allowed values of k. As in Sec. 2.1, these are determined by imposing the PBC conditions $\phi_{N+1} = \phi_1$ in Eq. (2.22). This leads to $e^{ikaN} = 1$ so that $k = 2\pi\ell/(Na)$, where $\ell = 0, \pm1, \pm2, \ldots$. Unlike the free particle case, however, k cannot take on an infinite number of values. The Hilbert space in this case has dimension N, so we only need a set of N eigenvectors. What we need to

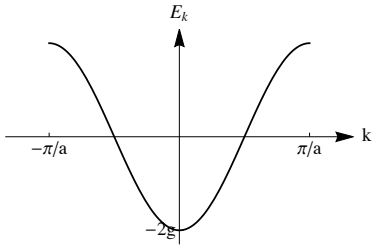

Figure 4: Dispersion relation (2.26) for the tight-binding model. The values of k are actually discretized as in Eq. (2.27) and restricted to the first Brillouin zone $k \in [-\pi/a, \pi/a]$.

look for are the eigenvectors. The eigenvalues can be equal (that is just a degeneracy). But the eigenvectors have to be distinct to form a orthonormal basis. In (2.22) the quantities n are integers; thus, if we displace

$$k \to k + 2\pi/a$$

we do not change ϕ_n^k. As a consequence, it suffices to pick out only those values of k which are within an interval of length $2\pi/a$. The customary choice is a symmetric interval

$$k \in [-\frac{\pi}{a}, \frac{\pi}{a}], \tag{2.24}$$

known historically as the **first Brillouin zone**. The correct choice for the values of k are therefore[2]

$$k = \frac{2\pi\ell}{Na}, \qquad -\frac{N}{2} < \ell \leq \frac{N}{2}.$$

To summarize, the eigenvectors of the tight-binding Hamiltonian (2.17) are plane waves of the form

$$\phi_n^k = \frac{e^{ikx_n}}{\sqrt{N}}, \tag{2.25}$$

whereas the dispersion relation is

$$E_k = -2g\cos ka. \tag{2.26}$$

In both formulas, the quantum numbers k are discretrized as

$$k = \frac{2\pi\ell}{Na}, \qquad -\frac{N}{2} < \ell \leq \frac{N}{2}. \tag{2.27}$$

where ℓ is an integer. This yields values of k in the first Brillouin zone $k \in [-\pi/a, \pi/a]$. The dispersion relation is shown in Fig. 4.

[2] Notice how I wrote $<$ and $\leq$ in this expression. This is simply to ensure that we really take N distinct eigenvectors. If N is odd, like $N = 5$, then ℓ can take on 5 values $-2, -1, 0, 1, 2$. But if N is even, like $N = 6$, then we cannot take $-3, -2, -1, 0, 1, 2, 3$. That gives 7 values. It turns out that in this case -3 and 3 give the same eigenvector ϕ_n^k, so it suffices to choose one of them. I chose $+N/2$.

Tight-binding is kind of a free particle

The dispersion relation in Fig. 4 is, similarly to Fig. 1, not really a continuous function. Rather, it is discrete in small steps of $\Delta k = 2\pi/Na$. The big difference between the tight-binding and free particle dispersion relations is that the latter is unbounded, with k going all the way to $\pm\infty$. In the tight-binding, on the other hand, k is restricted to the first Brillouin zone $k \in [-\pi/a, \pi/a]$.

This introduces the notions of **ultra-violet** and **infrared** cutoffs, which are nicknames we give for stuff with high and low energy respectively. The fact that the tight-binding dispersion relation is restricted to $[-\pi/a, \pi/a]$ is an ultra-violet cutoff. And the fact that it is discrete, in steps of $\Delta k = 2\pi/Na$, is an infrared cutoff. Ultraviolet is related to the lattice spacing a, whereas infrared is related to the total size $L = Na$ of the chain. The free particle has the exact same infrared cutoff [Eq. (2.9)]. But no ultraviolet since it has no underlying discrete lattice.

We are always interested in the limit of large chain sizes N (or large lengths $L = Na$). Thus, the infrared cutoff is always vanishingly small. Sometimes we may also be interested in the limit where the lattice spacing a is very small. In this case we may expand Eq. (2.26) in a, leading to

$$E_k \simeq -2g + ga^2k^2. \tag{2.28}$$

The first term is a constant and energy is only defined up to a constant. What matters is the second term. If we define an **effective mass**

$$m_e = \frac{1}{2ga^2}, \tag{2.29}$$

we then find that

$$E_k \simeq \frac{k^2}{2m_e} \tag{2.30}$$

For small ka we therefore see that the tight-binding model predicts the same kind of non-relativistic dispersion relation as the free-particle, but with an effective mass m_e. In fact, from Eq. (2.29) we see that $m_e \propto 1/g$, so that high tunneling rates make the particles very light. This effective mass can actually be measured in semiconductors and can vary quite dramatically, from 0.01 to 20 times the electron mass. Thus, in effect, an electron can move through a crystal as if it were a free particle, but with a completely different mass. Pretty cool huh?

2.3 Aubry-Andre model

Let us consider once again the tight-binding model (2.17), but introduce a **lattice dependent potential** so that the Hamiltonian changes to

$$H = \sum_{n=1}^{N} V_n |n\rangle\langle n| - g \sum_{i=1}^{N} \Big\{ |n\rangle\langle n+1| + |n+1\rangle\langle n| \Big\}, \tag{2.31}$$

where V_n are arbitrary numbers. In matrix form this looks like (e.g. for $N = 5$):

$$H = \begin{pmatrix} V_1 & -g & 0 & 0 & -g \\ -g & V_2 & -g & 0 & 0 \\ 0 & -g & V_3 & -g & 0 \\ 0 & 0 & -g & V_4 & -g \\ -g & 0 & 0 & -g & V_5 \end{pmatrix}. \tag{2.32}$$

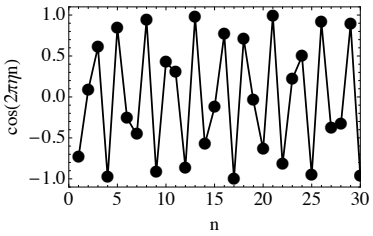

Figure 5: The function $\cos(2\pi\eta n)$, where $\eta = \frac{1}{2}(\sqrt{5}+1)$ is the Golden Ratio, is quasi-periodic and never repeats itself.

This Hamiltonian can still be treated in pretty much the same way. For instance, if we wish to find its eigenvalues and eigenvectors, we can generalize Eq. (2.21) to read

$$\boxed{V_n\phi_n - g(\phi_{n-1} + \phi_{n+1}) = E\phi_n.} \tag{2.33}$$

For general V_n it is quite unlikely we will be able to solve this equation analytically. But we can notwithstanding diagonalize it numerically. Just construct a big matrix like (2.32) and find its eigenvalues and eigenvectors. The task is not too bad because the matrix is tridiagonal (most entries are zero).

By playing with the potential V_n one can find an incredibly rich set of behaviors. *Really*. You will be amazed. One particularly nice choice, first studied by Aubry and Andre, is to choose the V_n to be of the form

$$V_n = \lambda\cos(2\pi\eta n), \tag{2.34}$$

where λ is a constant and $\eta = \frac{1}{2}(\sqrt{5}+1)$ is the Golden Ratio. The reason for choosing such a weird potential is because it is **quasi-periodic**: the irrationality of η makes it so that the V_n never repeat. The Hamiltonian (2.31) with the potential (2.34), is called the **Aubry-Andre model**.

This model presents a **localization transition**. A localization transition is not a phase transition. It is a transition on the **eigenvectors** of the model. The transition occurs at $\lambda = 1$. When $\lambda = 0$ the model is just the original tight-binding Hamiltonian (2.17), where the eigenvectors are plane waves [Eq. (2.22)]. Plane waves are *extended* in space; they are delocalized. We see this by analyzing $|\phi_n^k|^2$ as a function of the position n. For plane waves this would be just $|\phi_n^k|^2 = 1/N$ so all eigenvectors of the tight-binding model are completely flat in space; the probability of finding the particle is independent of the site n.

If we now start to change λ, we get some kind of spatial dependence. Since for $\lambda \neq 0$ the k are no longer good quantum numbers, I will use a generic index α to label the eigenstates. Plots for the ground-state probabilities $|\phi_n^1|^2$ as a function of n are shown in Fig. 6 for $N = 100$ sites and different values of λ. As can be seen, when $\lambda < 1$ the eigenvectors are extended through all of space. Even when $\lambda = 0.9$ or 0.95, the eigenvectors are extremely spread out. But when we reach $\lambda = 1$, sorcery happens: the eigenvalues become extremely localized in a specific region of space. The amount it localizes depends on the number of sites N. For large N the localization becomes sharper and sharper. Moreover, the position where it localizes is not at all obvious, but you can tune it if you add a phase to Eq. (2.34), as $V_n = \lambda\cos(2\pi\eta n + \theta)$. For $\lambda > 1$ the eigenvectors remain localized.

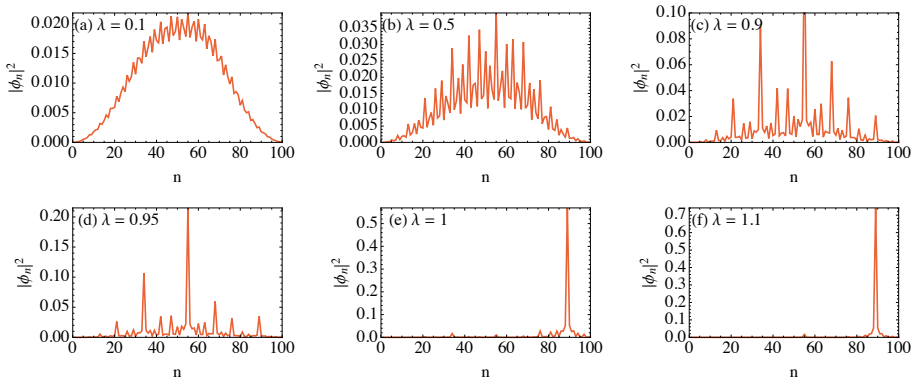

Figure 6: The ground-state wavefunction $|\phi_n|^2$ for the Aubry-Andre model (2.34) for different values of λ, with $g = 1$ and $N = 100$.

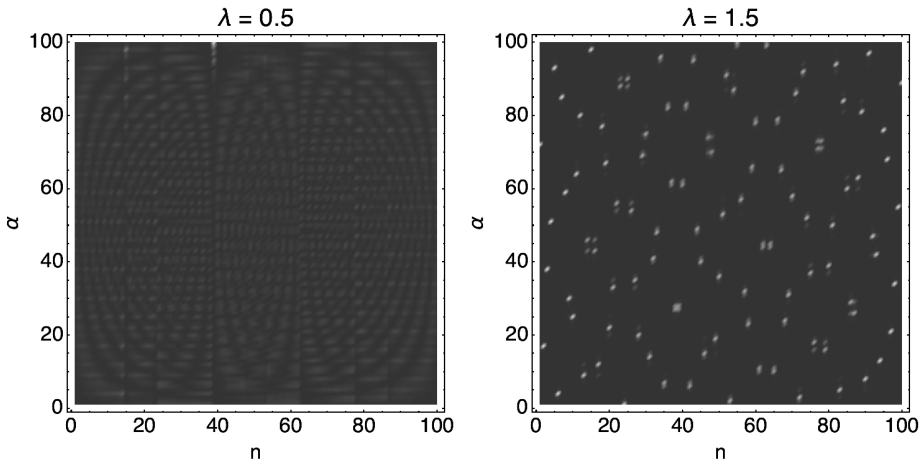

Figure 7: $|\phi_n^\alpha|^2$ for all eigenvectors of the Aubry-Andre model for $\lambda = 0.5$ (left) and $\lambda = 1.5$ (right), with $g = 1$ and $N = 100$. Here α is just an index labeling the eigenvectors.

The localization transition for the Aubry-Andre model does not occur only for the ground-state. *It occurs for the entire spectrum.* This is shown in Fig. 7 where I plot all eigenvectors for a value of λ below and above the transition. The appearance of spots in the graph when $\lambda > 1$ shows clearly that all eigenvectors are localized.

A more mathematical way of **quantifying localization** is through the Inverse Participation Ratio (IPR). For an arbitrary state $|\psi\rangle = \sum_n \psi_n |n\rangle$ the IPR is defined as

$$\text{IPR}(\psi) = \sum_n |\psi_n|^4. \tag{2.35}$$

When the state is completely localized the IPR will be 1 and when it is completely delocalized [like the plane waves (2.22)] it becomes

$$\text{IPR}(\psi) = \frac{1}{N^2}\sum_{n=1}^{N} 1 = \frac{1}{N}.$$

For large sizes, this tends to zero. Thus, the IPR is a number which varies essentially

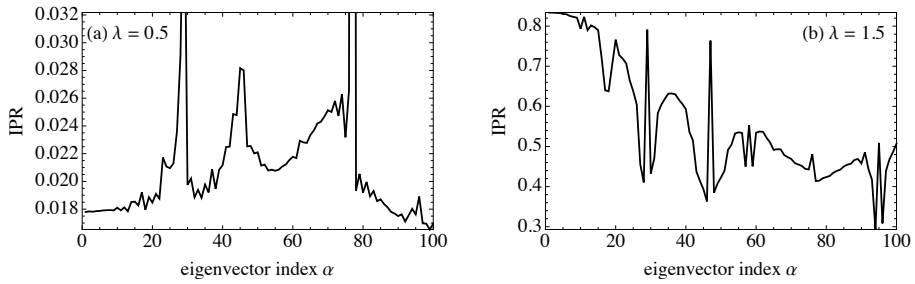

Figure 8: Inverse Participation Ratio (IPR) [Eq. (2.35)] for all eigenvectors of the Aubry-Andre modelfor $\lambda = 0.5$ (left) and $\lambda = 1.5$ (right), with $g = 1$ and $N = 100$.

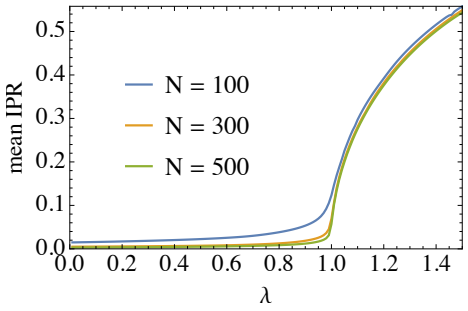

Figure 9: Mean IPR as a function of λ for different values of N and $g = 1$.

between 0 and 1. The larger is the IPR, the more localized is a state. In Fig. 8 I plot the IPR for the different eigenvectors of the Aubry-Andre model, below and above the transition. The plots are quite irregular. But notice the vertical scale. It is tiny below the transition (extended phase) and of order 1 above it (localized phase). This becomes even clearer if we plot the mean IPR (averaged over all eigenvectors), as shown in Fig. 9. The plot clearly shows a transition occurring at $\lambda = 1$, which becomes sharper and sharper as $N \to \infty$.

The Aubry-Andre model may at first seem simple, but it is not. Doing anything analytically with this model is notoriously difficult and there is much about it (and related models) which is still unknown. The Brazilian mathematician, Prof. Artur Ávila, gave seminal contributions to this field for instance, which was part of the reason he was awarded the Fields Medal. You can read about it, for instance, in arXiv:math/0503363.

3 Creation and annihilation operators

We are now finally ready to introduce the idea of second quantization. In quantum mechanics the number of particles is fixed. We may work with 1 particle. Or with 2 particles. Or with 3. When say something like "consider a system of N spins", the number of particles was fixed at N. The jump to second quantization is quite simple: we lift this constraint and **assume that the number of particles may fluctuate**. To describe this, we introduce creation and annihilation operators, similarly to what we do in the case of the harmonic oscillator, but which **create or annihilate actual particles**. This is the reason why we need single-particle states: these operators create and

annihilate particles *in* single-particle states.

Here is how it works. Let $|\alpha\rangle$ denote an arbitrary set of orthonormal single-particle states satisfying

$$\langle\alpha|\beta\rangle = \delta_{\alpha,\beta}. \tag{3.1}$$

We then define an operator $a_\alpha^\dagger$ and a special state $|0\rangle$, called the **vacuum**, such that

$$a_\alpha^\dagger|0\rangle = |\alpha\rangle \equiv |1_\alpha\rangle. \tag{3.2}$$

The operator $a_\alpha^\dagger$ therefore creates a particle in the single-particle state $|\alpha\rangle$. The notation $|1_\alpha\rangle$ is merely to emphasize that this is a state with one particle in state $|\alpha\rangle$. Similarly, the annihilation operator a_α is designed to annihilate a particle in the state $|\alpha\rangle$:

$$a_\alpha|\beta\rangle = \delta_{\alpha,\beta}|0\rangle. \tag{3.3}$$

If $\beta \neq \alpha$ then a_α cannot annihilate and we get zero. But if $\beta = \alpha$ then a_α destroys a particle and we are back to the vacuum.

3.1 Bosons and Fermions

Things become more interesting when we create multiple particles. Second quantization is specifically designed to describe **identical particles**. In 1940 Wolfgang Pauli published a paper entitled *"The connection between spin and statistics"* where he shows that, as a consequence of the Lorentz group of special relativity, identical particles can behave in one of two ways. **Bosons** are symmetric with respect to creation of two particles:

$$a_\beta^\dagger a_\alpha^\dagger|0\rangle = a_\alpha^\dagger a_\beta^\dagger|0\rangle := |1_\beta, 1_\alpha\rangle, \tag{3.4}$$

whereas **Fermions** are anti-symmetric:

$$a_\beta^\dagger a_\alpha^\dagger|0\rangle = -a_\alpha^\dagger a_\beta^\dagger|0\rangle := |1_\beta, 1_\alpha\rangle. \tag{3.5}$$

In the case of Bosons, due to (3.4), we can move the states around at will: $|1_\beta, 1_\alpha\rangle = |1_\alpha, 1_\beta\rangle$. For Fermions, however, every time we move the states around we get a minus sign: $|1_\beta, 1_\alpha\rangle = -|1_\alpha, 1_\beta\rangle$.

Pauli also showed that Bosons and Fermions have different spin values: Bosons' spins are integer valued (0,1,2,...) whereas Fermions' spins are half-integers ($1/2, 3/2, \ldots$). In the case of Fermions, in particular, if we set $\beta = \alpha$ in Eq. (3.5) we see that

$$(a_\alpha^\dagger)^2|0\rangle = -(a_\alpha^\dagger)^2|0\rangle, \tag{3.6}$$

which implies

$$\boxed{(a_\alpha^\dagger)^2 = 0 \qquad \text{(for Fermions).}} \tag{3.7}$$

This is the **Pauli exclusion principle**: it is forbidden to create two fermions on the same single-particle state $|\alpha\rangle$.

Eq. (3.4) implies that bosonic creation and annihilation operators should **commute**:

$$[a_\alpha, a_\beta] = [a_\alpha^\dagger, a_\beta^\dagger] = 0, \tag{3.8}$$

whereas Eq. (3.5) implies that fermionic operators should **anti-commute**:

$$\{a_\alpha, a_\beta\} = \{a_\alpha^\dagger, a_\beta^\dagger\} = 0, \tag{3.9}$$

where $\{A, B\} = AB + BA$ is the anti-commutator.

In addition, one can also show that similar properties hold for the mixed commutation relations between a_α and $a_\beta^\dagger$. In the case of Bosons, we get the typical harmonic-oscillator like commutation relations

$$[a_\alpha, a_\beta^\dagger] = \delta_{\alpha,\beta}. \tag{3.10}$$

whereas in the case of Fermions we get the same thing, but with anti-commutators:

$$\{a_\alpha, a_\beta^\dagger\} = \delta_{\alpha,\beta}. \tag{3.11}$$

These algebraic relations are the building blocks of second quantization.

Fock space for Bosons

Fock space is the generalization of Hilbert space to the case where the number of particles is not fixed. In the case of Bosons the operators a_α behave exactly like harmonic oscillator operators (see Appendix A.2). So first there is a state $|0\rangle$ with no particles. Then there is a set of states $|\alpha\rangle \equiv |1_\alpha\rangle$ with exactly one particle. Next there is a set of states with two particles, which include $|1_\beta, 1_\alpha\rangle$ as well as states of the form $|2_\alpha\rangle$. More generally, the Fock states of a Bosonic system have the form

$$|n_1, n_2, n_3, \ldots\rangle, \qquad n_\alpha = 0, 1, 2, \ldots = \text{ number of particles in state } |\alpha\rangle.$$

Here I am labeling the single-particle states as $\alpha = 1, 2, 3, \ldots$. But, of course, you can label them anyway you want. For instance, if the Fock states are the position states $|i\rangle$, $i = 1, 2, \ldots, N$ for a **tight-binding lattice**, then the Fock states will have the form $|n_1, n_2, \ldots, n_N\rangle$ where n_{42} is the number of particles in site 42. Bosons don't satisfy the Pauli exclusion principle, so we can have 20 particles on site 42, or whatever. Notice also how in this tight-binding example, the number of single-particle states is finite (there are N of them), but the number of Fock states is infinite, since each site can have any number of particles.

We can characterize the number of particles by defining the **number operator**

$$\hat{N} = \sum_\alpha a_\alpha^\dagger a_\alpha. \tag{3.12}$$

I usually don't put hats on operators; but in this case it is useful since the letter N appears far too often in physics! The number operator acting on the states above then yield

$$\hat{N}|0\rangle = 0,$$

$$\hat{N}|1_\alpha\rangle = |1_\alpha\rangle,$$

$$\hat{N}|1_\beta, 1_\alpha\rangle = 2|1_\beta, 1_\alpha\rangle,$$

or, more generally,

$$\hat{\mathcal{N}}\,|n_1, n_2, n_3, \ldots\rangle = \Big(\sum_\alpha n_\alpha\Big)|n_1, n_2, n_3, \ldots\rangle, \tag{3.13}$$

where the sum here is over all allowed single-particle states.

Using the eigenvalues of $\hat{\mathcal{N}}$ we can factor the Fock space into sectors with different numbers of particles. We can also move from one sector to the other. This is where a_α and $a_\alpha^\dagger$ come in. Because of the commutation relations (3.10), the rules for applying a_α and $a_\alpha^\dagger$ are exactly like those of the harmonic oscillator:

$$a_\alpha^\dagger|n_1, \ldots, n_\alpha, \ldots\rangle = \sqrt{n_\alpha + 1}\;|n_1, \ldots, n_\alpha + 1, \ldots\rangle, \tag{3.14}$$

$$a_\alpha|n_1, \ldots, n_\alpha, \ldots\rangle = \sqrt{n_\alpha}\;|n_1, \ldots, n_\alpha - 1, \ldots\rangle. \tag{3.15}$$

Fock space for Fermions

The situation for Fermions is almost identical. The ony difference is that Fermions satisfy the Pauli exclusion principle (3.7). This means that we can never put more than two particles in the same single-particle state (see Appendix A.3). The Fock space for Fermions will therefore have the form

$$\boxed{|n_1, n_2, n_3, \ldots\rangle, \qquad n_\alpha = 0, 1 = \text{ number of particles in state } |\alpha\rangle.}$$

Looks exactly like the bosonic case, but $n_\alpha = 0, 1$. Fermionic Fock states also have the ambiguity we saw before, about which particle we create first: $|1_\beta, 1_\alpha\rangle = -|1_\alpha, 1_\beta\rangle$. This ambiguity has no physical consequences, as long as you are consistent with your definitions.

3.2 Transformations between creation and annihilation operators

Creation operators transform like kets

Consider now two sets of single-particle states $\{|\alpha\rangle\}$ and $\{|i\rangle\}$. To each set we can attribute corresponding creation and annihilation operators a_α and b_i. How are these two sets of operators connected? That is, how to transform from a_α to b_i? The bases $\{|\alpha\rangle\}$ and $\{|i\rangle\}$ can be connected by a unitary transformation

$$|\alpha\rangle = \sum_i |i\rangle\langle i|\alpha\rangle = \sum_i U_{\alpha i}|i\rangle, \tag{3.16}$$

where $U_{\alpha i} = \langle i|\alpha\rangle$. Now remember Eq. (3.2): the creation operators are defined by the relations

$$|\alpha\rangle = a_\alpha^\dagger|0\rangle, \tag{3.17}$$

$$|i\rangle = b_i^\dagger|0\rangle \tag{3.18}$$

Using Eq. (3.16) we then get

$$a_\alpha^\dagger|0\rangle = |\alpha\rangle = \sum_i U_{\alpha i}|i\rangle = \sum_i U_{\alpha i} b_i^\dagger|0\rangle.$$

Whence,

$$\boxed{a_\alpha^\dagger = \sum_i \langle i|\alpha\rangle b_i^\dagger.} \tag{3.19}$$

This is exactly like the transformation rule (3.16). Creation operators transform like kets! :)

Eq. (3.16) connects the two sets a_α and b_i by means of a unitary transformation $U_{\alpha,i} = \langle i|\alpha\rangle$. A very important property of second quantization is that **unitary transformations preserve the algebra**. Let us check this. We start with Bosons and Eq. (3.10): assume that $[b_i, b_j^\dagger] = \delta_{ij}$. Then

$$[a_\alpha, a_\beta^\dagger] = \sum_{i,j} U_{\alpha i}^* U_{\beta j}\, [b_i, b_j^\dagger] = \sum_{i,j} U_{\alpha i}^* U_{\beta j}\delta_{ij} = \sum_i U_{\alpha i}^* U_{\beta i}.$$

The resulting sum is nothing but

$$\sum_i U_{\alpha i}^* U_{\beta i} = \sum_i U_{\beta i}(U^\dagger)_{i\alpha} = (UU^\dagger)_{\beta,\alpha} = \delta_{\alpha,\beta},$$

since U is assumed to be unitary. Thus we conclude that

$$\boxed{[b_i, b_j^\dagger] = \delta_{ij} \quad \rightarrow \quad [a_\alpha, a_\beta^\dagger] = \delta_{\alpha\beta}.} \tag{3.20}$$

Unitary transformations preserve the algebra. I will leave it for you as an exercise to show that this is also true for Fermions. Please do it. It's almost identical as the calculation above, but it is a good exercise.

Number operator

If we choose a basis set $\{|\alpha\rangle\}$ as our single-particle states, then the number operator is $\hat{\mathcal{N}} = \sum_\alpha a_\alpha^\dagger a_\alpha$. But if we use a basis set $\{|i\rangle\}$ the number operator would be $\hat{\mathcal{N}} = \sum_i b_i^\dagger b_i$. Given that the number operator *counts* the number of particles in the system, it would be somewhat weird if $\hat{\mathcal{N}}$ depended on *our* choice of basis. Indeed, it does not. We can check this using again the unitarity of the transformation (3.19):

$$\hat{\mathcal{N}} = \sum_\alpha a_\alpha^\dagger a_\alpha = \sum_{\alpha,i,j} U_{\alpha i} U_{\alpha j}^* b_i^\dagger b_j.$$

We now carry first the sum over α, which yields a δ_{ij}. Thus

$$\hat{\mathcal{N}} = \sum_\alpha a_\alpha^\dagger a_\alpha = \sum_i b_i^\dagger b_i. \tag{3.21}$$

The number operator is independent of the choice of basis.

Commutation relations

We can also use these results to establish the commutation relations between a_α and b_i. Again I will focus on Bosons, with the case of Fermions being identical. Using Eq. (3.19) we find

$$[a_\alpha, b_i] = \sum_j \langle j|\alpha\rangle^* [b_j, b_i^\dagger].$$

But since $[b_j, b_i^\dagger] = \delta_{ij}$ this reduces to

$$[a_\alpha, b_i^\dagger] = \langle \alpha | i \rangle. \tag{3.22}$$

The relation for Fermions is identical, but with the anti-commutator. This result is quite nice, as it shows how to generalize the idea of creation and annihilation operators to arbitrary single-particle states (which do not necessarily form a basis).

In fact, from now on we can generalize our notations a bit and write, instead of Eqs. (3.10) and (3.11),

$$[a_\alpha, a_\beta^\dagger] = \langle \alpha | \beta \rangle, \qquad \text{(Bosons)} \tag{3.23}$$

$$\{a_\alpha, a_\beta^\dagger\} = \langle \alpha | \beta \rangle, \qquad \text{(Fermions)} \tag{3.24}$$

where $|\alpha\rangle$ and $|\beta\rangle$ can now be arbitrary single-particle kets (not necessarily orthogonal). If they happen to be orthogonal, we of course recover what we had before. The important message conveyed by these formulas is that *the amount by which a_α and $a_\beta^\dagger$ fail to commute is related to the overlap of their single-particle wavefunctions.*

3.3 Non-interacting Hamiltonians in the language of second quantization

So far we have only dealt with the algebraic structure of second quantization. In order to make this useful, we have to start expressing Hamiltonians in this new language. This turns out to be quite intuitive. Let's work through some examples and you will see.

Free particle

Consider the free particle in Sec. 2.1. There, we saw that a set of single-particle states were the momentum states $|k\rangle$, where $k = 2\pi\ell/L$ and $\ell = 0, \pm 1, \pm 2, \ldots$. We may thus work with operators a_k and $a_k^\dagger$ which annihilate and create particles with momentum k. Now consider again the single-particle Hamiltonian for the free particle:

$$H_1 = \frac{p^2}{2m}, \tag{3.25}$$

where I am calling this H_1 here just to emphasize that this is a single particle Hamiltonian. Suppose now that we have N **non-interacting** free particles. The N particle Hamiltonian would be

$$H_N = \sum_{i=1}^{N} \frac{p_i^2}{2m}.$$

But if we don't want to specify what is the number of particles, then this expression doesn't really work.

Second quantization approaches the problem differently (and much more elegantly). In second quantization the Hamiltonian of non-interacting free particles is expressed as

$$\mathcal{H} = \sum_k \frac{k^2}{2m} a_k^\dagger a_k. \tag{3.26}$$

The logic behind this formula is super simple: $a_k^\dagger a_k$ *counts* how many particles have momentum k. So we sum over all k with the energy levels $k^2/2m$ *weighted* by the number of particles with that k value. The Hamiltonian (3.26) does not live in Hilbert space. It lives in Fock space; in the space where the number of particles is not fixed. Of course, the same logic also applies for other operators. For instance, we can define the momentum operator in second quantization as

$$\mathcal{P} = \sum_k k\, a_k^\dagger a_k. \tag{3.27}$$

This operators quantifies the total momentum contained in the system, irrespective of how many particles are there. Compare this also with the number operator:

$$\hat{\mathcal{N}} = \sum_k a_k^\dagger a_k.$$

The idea is similar, except that in $\mathcal{P}$ each $a_k^\dagger a_k$ is weighted by k.

Non-interacting systems

Eq. (3.26) is an example of a non-interacting Hamiltonian. We can have any number of particles (identical, of course), but they do not directly interact. We will learn how to deal with interactions in a second. But before doing so, let us try to formalize a bit better how to write down second-quantized versions of non-interacting Hamiltonians. It turns out there is a super easy way to do this.

Consider a single-particle Hamiltonian H_1 and suppose it can be diagonalized as

$$H_1 = \sum_\alpha \varepsilon_\alpha |\alpha\rangle\langle\alpha|. \tag{3.28}$$

The eigenvectors $|\alpha\rangle$ of H_1 of course form a basis of single-particle states. Thus, we can define operators a_α. The second-quantized version of (3.28) is then simply

$$\mathcal{H} = \sum_\alpha \varepsilon_\alpha\, a_\alpha^\dagger a_\alpha. \tag{3.29}$$

Notice the similarity in structure between (3.28) and (3.29). We essentially replace $|\alpha\rangle\langle\alpha|$ with $a_\alpha^\dagger a_\alpha$. The logic is also the same as before: $a_\alpha^\dagger a_\alpha$ counts how many particles are in eigenstate $|\alpha\rangle$. The Hamiltonian is then the sum of all possible energies ε_α weighted by this number of particles.

What about the situation where H_1 is not diagonal? Suppose we are given a generic

single-particle Hamiltonian of the form

$$H_1 = \sum_{i,j} H_{i,j}|i\rangle\langle j|, \tag{3.30}$$

where $H_{i,j} = \langle i|H_1|j\rangle$. What to do? We first diagonalize (3.30) by introducing a unitary transformation:

$$|i\rangle = \sum_{\alpha} |\alpha\rangle\langle\alpha|i\rangle. \tag{3.31}$$

This turns (3.30) into (3.28). The corresponding second-quantized Hamiltonian is then given by Eq. (3.29). Now we use the transformation rule (3.19) for creation operators:

$$\mathcal{H} = \sum_{\alpha} \varepsilon_\alpha a_\alpha^\dagger a_\alpha = \sum_{\alpha,i,j} \varepsilon_\alpha \langle i|i\alpha\rangle\langle\alpha|j\rangle a_i^\dagger a_j. \tag{3.32}$$

We can rearrange this as:

$$\mathcal{H} = \sum_{i,j} \left(\sum_{\alpha} \langle i|\alpha\rangle \varepsilon_\alpha \langle\alpha|j\rangle \right) a_i^\dagger a_j.$$

The thing inside parenthesis is

$$\sum_{\alpha} \langle i|\alpha\rangle \varepsilon_\alpha \langle\alpha|j\rangle = \langle i|H_1|j\rangle = H_{ij},$$

which are coefficients in Eq. (3.30). Thus, we can write (3.32) as

$$\mathcal{H} = \sum_{i,j} H_{ij} a_i^\dagger a_j.$$

To summarize, we thus have that the single-particle Hamiltonian

$$H_1 = \sum_{ij} H_{ij}|i\rangle\langle j|, \tag{3.33}$$

is written in the language of second-quantization as

$$\mathcal{H} = \sum_{ij} H_{ij} a_i^\dagger a_j. \tag{3.34}$$

Again, notice the similarity: we simply have to replace kets $|i\rangle$ with $a_i^\dagger$ and $\langle j|$ with a_j. This important conclusion is also readily generalized to other operators. Let A_1 denote *any* single-particle operator and $\{|i\rangle\}$ any set of single-particle states. Then, the corresponding second-quantized version of A_1 will be

$$\mathcal{A} = \sum_{i,j} A_{ij} a_i^\dagger a_j, \qquad A_{ij} = \langle i|A_1|j\rangle, \tag{3.35}$$

irrespective of whether or not A_1 is diagonal in this basis.

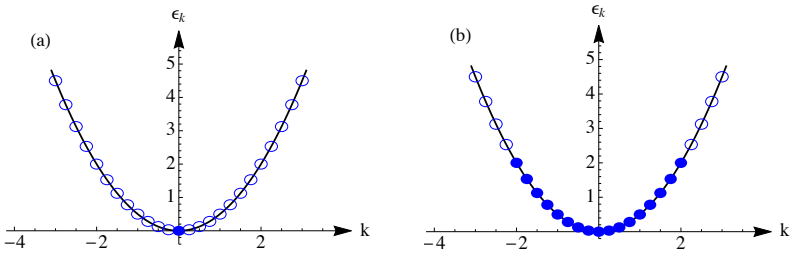

Figure 10: Ground-state of N free particles in 1D for the case of (a) Bosons and (b) Fermions.

Ground state of a free particle system

Let us return to the free particle Hamiltonian (3.26). What is the ground-state? Well, this is a tricky question, because it depends on whether we assume the number of particles is fixed or not. Each term in Eq. (3.26) is strictly positive, so the more particles we add, the higher will the energy be. If the number of particles can vary then, of course, the ground-state will be that state with zero particles.

But suppose that the number of particles is fixed at some value N. Then how to construct the state with the smallest possible energy? Well, here we have to distinguish between Bosons and Fermions. Bosons do not satisfy the exclusion principle, so we can put as many of them as we want in the same state. In Eq. (3.26) the single-particle ket with the lowest energy is that with $k = 0$. Thus, for the case of Bosons the ground-state will be a state where $k = 0$ is populated with N particles; viz.,

$$\boxed{|\Psi_{\rm gs}\rangle = \frac{(a_0^\dagger)^N}{\sqrt{N!}}|0\rangle,} \tag{3.36}$$

where $a_0^\dagger = a_{k=0}^\dagger$ is the creation operator at momentum $k = 0$ (see Fig. 10(a)). The factor of $\sqrt{N!}$ is put here simply for normalization [see Appendix (A.2) and Eq. (A.12) for more details]. The state (3.36) is roughly what happens in a Bose-Einstein condensate (which we will study later on). Bose-Einstein condensation (BEC) is not a condensation in position space; it is a condensation in momentum space. What happens in the BEC is that a macroscopically large number of particles tend to settled down in the same lowest energy state, which is precisely the state (3.36).

Next let us analyze the ground-state of the free particle system in the case of Fermions. Now there is the Pauli exclusion principle to care about. Let us assume, first, that there is no spin. Then the state $k = 0$ can hold only one Fermion. Once it is filled, we have no choice but to put the second particle in the next lowest energy state. In this case, remembering how the k are quantized as in Eq. (2.9), this state will be $k = \pm 2\pi/L$. We then keep on going, filling out successive k states of increasing E_k until we fill the N particles. The situation will then be like in Fig. 10(b).

The largest filled value of k is called the **Fermi momentum** k_F. It depends on the value of N. We will learn more sophisticated methods of finding k_F later on. For the 1D case considered here, if we forget about the state with $k = 0$, then essentially each value of $|k|$ will take on two particles. So given that $\Delta k = 2\pi/L$, we get In fact, in the 1D case we simply have

$$k_F = \frac{\pi N}{L}. \tag{3.37}$$

The ground-state of a Fermionic system can then be written as

$$\boxed{|\Psi_{\rm gs}\rangle = \Big(\prod_{|k|\leq k_F} a_k^\dagger\Big)|0\rangle.} \tag{3.38}$$

Spelled out in words: to construct the ground-state we create particles (by applying $a_k^\dagger$) on all k states having $|k| < k_F$, leaving the higher order momenta entirely empty. The number of particles N is contained in k_F.

The highest occupied energy is called the **Fermi energy** ε_F (everything with the name Fermi in this context refers to "highest occupied..."). Since $\varepsilon_k = k^2/2m$, the Fermi energy will be

$$\varepsilon_F = \frac{k_F^2}{2m} \tag{3.39}$$

The Fermi energy is not the ground-state energy. Instead, the ground-state energy is

$$E_{\rm gs} = \sum_{|k|\leq k_F} \frac{k^2}{2m}. \tag{3.40}$$

Computing this sum if we have 42 particles can be annoying. But if N, as well as the length of the box L, are both large, the sum can be converted to an integral. We will come back to this below. For now I will leave $E_{\rm gs}$ like this.

Free particles with spin

If we have spin, then the single-particle kets should be replaced with $|k,\sigma\rangle$, so the annihilation operators become $a_{k\sigma}$. Let us focus on the case of Fermions. And let us first assume that there are no magnetic fields present (nor any kinds of interactions). The Hamiltonian of the system then becomes

$$\mathcal{H} = \sum_{k,\sigma} \frac{k^2}{2m} a_{k\sigma}^\dagger a_{k\sigma}. \tag{3.41}$$

The sum is now over all values of k and both spin components $\sigma = \pm 1$. The energies, however, do not depend on the spin value, but only on k (because we did not put a magnetic field yet). To construct the fermionic ground-state, we proceed exactly as before. The only difference is that now each k value can take on two fermions so the Fermi momentum k_F will be half of what we had before The ground-state can then be built as

$$|\Psi_{\rm gs}\rangle = \Big(\prod_{|k|\leq k_F,\sigma} a_{k\sigma}^\dagger\Big)|0\rangle. \tag{3.42}$$

Next let us introduce a magnetic field. As we know, this shifts the energy levels by $-h\sigma$ so that the Hamiltonian becomes

$$\mathcal{H} = \sum_{k,\sigma} \Big(\frac{k^2}{2m} - h\sigma\Big) a_{k\sigma}^\dagger a_{k\sigma}. \tag{3.43}$$

The energies of the system now split into two **energy bands**, corresponding to $\sigma = 1$ and $\sigma = -1$. To construct the ground-state we always need to remember the obvious:

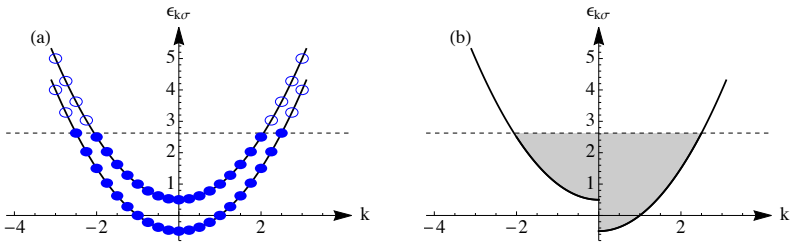

Figure 11: (a) Pictorial construction of the ground-state of the Hamiltonian (3.43) for the case of Fermions. (b) The energy bands in these kinds of problems are often represented in this form, with the band with $\sigma = +1$ plotted only for $k > 0$ and that with $\sigma = -1$ for $k < 0$. This does not mean that it is not possible to have $\sigma = +1$ and $k < 0$. It is simply for presentation purposes.

the ground-state is the state with smallest possible energy. The situation will then be somewhat like that in Fig. 11. There will be two bands, each with a discrete set of states. The states are then populated starting from the bottom, so we fill some states in one band and some states in the other band.

This kind of system is often represented pictorially as in Fig. 11(b), where the band with $\sigma = +1$ is plotted only for $k > 0$ and that with $\sigma = -1$ only for $k < 0$. This kind of plot is meant to emphasize the *population mismatch* between the spin up and spin down bands: because of the magnetic field there are now more spins in one band than the other. We can define the **magnetization operator** as

$$\mathcal{M} = \sum_{k,\sigma} \sigma \, a^\dagger_{k\sigma} a_{k\sigma} = \mathcal{N}_+ - \mathcal{N}_-, \tag{3.44}$$

where

$$\mathcal{N}_\sigma = \sum_k a^\dagger_{k\sigma} a_{k\sigma}, \tag{3.45}$$

are the populations of the two bands. The magnetization operator is thus the imbalance between the populations of the two bands.

The tight-binding model in the language of second quantization

An example of the above results is the tight-binding model (2.17). Using the recipe in Eq. (3.35), the corresponding Hamiltonian in the language of second quantization is

$$\boxed{\mathcal{H} = -g \sum_{i=1}^{N} \left(a^\dagger_i a_{i+1} + a^\dagger_{i+1} a_i \right).} \tag{3.46}$$

I personally think that written in this way the tight-binding Hamiltonian makes even more sense than Eq. (2.17). In fact, it gives a quite nice interpretation for what quantum tunneling *is*: tunneling means you annihilate a particle at site i and then create a particle at site $i + 1$.

Notice also how, even though $\mathcal{H}$ is creating particles in one place and annihilating in others, the process is such that particles cannot appear out of nowhere or go missing. In other words, **the Hamiltonian (3.46) preserves the number of particles**. This

is the nice thing about second quantization: we have a theory where the number of particles can fluctuate; the physical constraint that the number of particles is conserved is only implemented at the Hamiltonian. Since the Hamiltonian is the generator of the dynamics, if we start with 42 particles, as time moves on we will continue to have 42 particles. This can be stated mathematically as

$$[\mathcal{H}, \hat{N}] = 0, \tag{3.47}$$

where, in the case of the tight-binding Hamiltonian (3.35),

$$\hat{N} = \sum_{i=1}^{N} a_i^\dagger a_i. \tag{3.48}$$

I will leave for you to check that (3.47) is indeed satisfied. In fact, it suffices to check only that

$$[\hat{N}, a_i^\dagger a_j] = 0, \qquad \forall i, j. \tag{3.49}$$

I will leave this for you as an exercise.

3.4 Interacting Hamiltonians

Bose-Hubbard model

We have seen that single-particle Hamiltonians H_1 lead to second-quantized Hamiltonians which are **quadratic** in creation and annihilation operators. Quadratic Hamiltonians are thus a synonym with non-interacting systems. Interactions, on the other hand, are represented by Hamiltonians which are cubic or quartic. To illustrate the idea, let us consider a specific model called the **Bose-Hubbard** model. It is a tight-binding model for bosons, with interactions. The Hamiltonian is written as

$$\boxed{\mathcal{H} = -g \sum_{\langle i,j \rangle} (a_i^\dagger a_j + a_j^\dagger a_i) + \frac{U}{2} \sum_i a_i^\dagger a_i^\dagger a_i a_i.} \tag{3.50}$$

The first term is just tight-binding, which I slightly generalized to an arbitrary lattice by including a sum over nearest neighbors $\langle i, j \rangle$. The new thing is the last term. It describes a type of **contact interaction**. It is an "interaction" because it is quartic and it is "contact" because it acts on the same site.

To understand the logic behind this term, we use the bosonic algebra $a_i^\dagger a_i = a_i a_i^\dagger - 1$ to write

$$a_i^\dagger a_i^\dagger a_i a_i = a_i^\dagger (a_i a_i^\dagger - 1) a_i = a_i^\dagger a_i a_i^\dagger a_i - a_i^\dagger a_i.$$

We see here the appearance of the number operator $\hat{n}_i = a_i^\dagger a_i$:

$$a_i^\dagger a_i^\dagger a_i a_i = n_i (n_i - 1). \tag{3.51}$$

The interaction term therefore depends on the number of particles in each site. In table 1 I present a little table with the values of the energy $\frac{U}{2} n(n-1)$ for different values of n. We see that if there are 0 or 1 particles, the interaction term is zero. This is good. It means that there is *no self-interaction*. The last term in (3.50) only begins

Table 1: Interaction energies appearing in the Bose-Hubbard model (3.50).

n	0	1	2	3	4	5
$\frac{U}{2}n(n-1)$	0	0	U	$3U$	$6U$	$10U$

to contribute when there are 2 or more particles. In fact, $n(n-1)/2$ is precisely the number of ways for n particles to interact in pairs:

$$\binom{n}{2} = \frac{n!}{2!(n-2)!} = \frac{n(n-1)}{2}.$$

The Bose-Hubbard model is an important model in physics. It cannot be solved analytically, but there are tons of approximations which work quite well. Decades ago this model was just a theoretical exercise. But today it can be implemented in **ultra-cold atoms in optical lattices**,[3] thus making them one of the most interesting many-body models nowadays in physics.

Fermi-Hubbard model

More important than the Bose-Hubbard model is it's fermionic cousin, the Fermi-Hubbard model. It is again defined as a tight-binding lattice, but with an interaction of the form:

$$\mathcal{H} = -g \sum_{\langle i,j\rangle,\sigma} (a^\dagger_{i\sigma} a_{j\sigma} + a^\dagger_{j\sigma} a_{i\sigma}) + U \sum_i \hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow}. \tag{3.52}$$

Here I am using the notation $\sigma = \uparrow, \downarrow$ instead of $\sigma = \pm 1$, simply because this is how people usually write it. The logic behind this Hamiltonian is the following. Since Fermions satisfy the exclusion principle, we cannot have an interaction term like that of Eq. (3.50) because we cannot put two Fermions in the same state. What we *can* do, however, is put two Fermions in the same site, provided they have opposite spins. The interaction term is therefore dependent on $\hat{n}_{i,\uparrow}\hat{n}_{i,\downarrow}$ which will only give a non-zero contribution if both states $(i,\uparrow)$ and $(i,\downarrow)$ are filled.

It is widely believed that the Fermi-Hubbard model is capable of explaining the high temperature Cuprate superconductors that were discovered in the 80s. Solving this model would therefore explain what is, today, one of the most important unexplained phenomena in physics: while standard superconductivity can be explained by the so-called BCS theory, high-temperature superconductivity simply cannot. Explaining it could provide a guideline for devising superconductors at ambient temperature, which could revolutionize tons of applications. Needless to say, this model has no analytical solution. If you ever find it, book a plane ticket to Stockholm.

Electron-photon interactions

The Bose-Hubbard and Fermi-Hubbard Hamiltonians in Eqs. (3.50) and (3.52) contain an equal number of creation and annihilation operators. Consequently, they pre-

[3] See, for instance, M. Greiner, O. Mandel, T. Esslinger, T. W. Hänsch and I. Bloch, *"Quantum phase transition from a superfluid to a Mott insulator in a gas of ultracold atoms"*. Nature, **415** 39–44 (2002).

serve particle number. This is good because the particles here are electrons or atoms, so they cannot be destroyed or created out of nothing. Other particles, however, can. One example are the photons. Photonic Hamiltonians don't need to have the same number of creation and annihilation operators. If you shine light on a metal, for instance, the electrons may absorb the photon and gain some energy. Or they may emit a photon and slow down a bit.

Let c_k denote the fermionic annihilation operators for electrons and a_k the bosonic operators for the photons. The processes we just described can then be described by the following Hamiltonian

$$\mathcal{H} = \sum_{k,q} M_{kq}\Big(c^\dagger_{k+q} c_k a_q + c^\dagger_k c_{k+q} a^\dagger_q\Big). \tag{3.53}$$

The logic is the following: a term like $c^\dagger_{k+q} c_k a_q$ means that we annihilate a photon with momentum q to change the momentum of the electron from k to $k+q$. Of course, in second quantization we do that by first annihilating the electron with a_k and then recreating it with $a^\dagger_{k+q}$. The other term in Eq. (3.53) is, of course, just the adjoint of the first to make the Hamiltonian Hermitian.

The thing you should notice in the Hamiltonian (3.53) is that the number of electrons is conserved (as it should) but the number of photons is not. The physics of which particle should be conserved and which should not should thus always be encoded in the Hamiltonian.

4 Quadratic Hamiltonians

4.1 Recipe for diagonalizing quadratic Hamiltonians

Any Hamiltonian of the form (3.34), which is at most **quadratic** in creation and annihilation operators, represents **non-interacting systems**. As we will see below, interactions involve terms with 3 or more operators. Quadratic Hamiltonians can be readily diagonalized. In fact, we just saw how to do this by accident in the previous section. The goal of this section is to make these ideas clearer. In fact, we will develop a general recipe for diagonalizing these kinds of Hamiltonians which is extremely useful. The procedure is identical for Fermions and Bosons.

Suppose we start with a generic Hamiltonian of the form (3.34). Our goal is to put it in the form of Eq. (3.29). The first thing we need to do is to diagonalize the matrix of coefficients H_{ij}, which means going from Eq. (3.33) to Eq. (3.30). We do that by introducing the basis transformation (3.31), from a set of states $|i\rangle$ to a set of states $|\alpha\rangle$ which diagonalize H (from now on I will write H instead of H_1 to denote the single-particle Hamiltonian). This is equivalent to introducing a diagonal representation of the form

$$H = U\mathcal{E}U^\dagger, \qquad \text{or} \qquad H_{ij} = \sum_\alpha U_{i\alpha}\varepsilon_\alpha U^*_{j\alpha}$$

where $U_{i\alpha} = \langle i|\alpha\rangle$ is the basis transformation matrix and $\mathcal{E} = \text{diag}(\varepsilon_\alpha)$ is a diagonal matrix with the eigenenergies ε_α. Since $U_{i\alpha} = \langle i|\alpha\rangle$, the matrix U is a matrix whose *columns* contain the eigenvectors $|\alpha\rangle$.

Inserting this transformation in Eq. (3.34) we get

$$\mathcal{H} = \sum_{i,j,\alpha} U_{i\alpha}\varepsilon_\alpha U^*_{j\alpha} a^\dagger_i a_j.$$

We now regroup this sum as follows:

$$\mathcal{H} = \sum_{\alpha} \varepsilon_{\alpha} \Big(\sum_{i} U_{i\alpha} a_i^{\dagger} \Big) \Big(\sum -j U_{j\alpha}^{*} a_j \Big).$$

This motivates us to introduce a new set of operators

$$b_{\alpha}^{\dagger} = \sum_{i} U_{i\alpha} a_i^{\dagger}.$$

Remember: creation operators transform like kets. Since $U_{i\alpha}$ are unitary, the inverse transformation is simply:

$$a_i^{\dagger} = \sum_{\alpha} U_{i\alpha}^{*} b_{\alpha}^{\dagger}.$$

With this the Hamiltonian becomes

$$\mathcal{H} = \sum_{\alpha} \varepsilon_{\alpha} b_{\alpha}^{\dagger} b_{\alpha}.$$

which is what we set out to do.

Let me summarize this procedure: suppose we start with a 2nd-quantized Hamiltonian of the form

$$\mathcal{H} = \sum_{ij} H_{ij} a_i^{\dagger} a_j, \tag{4.1}$$

To diagonalize it, we first diagonalize the single-particle Hamiltonian H:

$$H = U \mathcal{E} U^{\dagger}, \qquad \text{or} \qquad H_{ij} = \sum_{\alpha} U_{i\alpha} \varepsilon_{\alpha} U_{j\alpha}^{*} \tag{4.2}$$

From this, introduce a new set of operators according to

$$b_{\alpha}^{\dagger} = \sum_{i} U_{i\alpha} a_i^{\dagger}, \qquad a_i^{\dagger} = \sum_{\alpha} U_{i\alpha}^{*} b_{\alpha}^{\dagger}. \tag{4.3}$$

Since $U_{i\alpha}$ is unitary, this preserves the algebra. Inserting these transformations, the Hamiltonian (4.1) becomes

$$\mathcal{H} = \sum_{\alpha} \varepsilon_{\alpha} b_{\alpha}^{\dagger} b_{\alpha}. \tag{4.4}$$

You may also wonder: why writing (4.1) in the form (4.4) means we *diagonalized* the Hamiltonian? We diagonalized it because now the Hamiltonian (4.4) is written as a sum of independent number operators $b_{\alpha}^{\dagger} b_{\alpha}$. We therefore know how to construct the corresponding **Fock space**. The eigenvalues of $\mathcal{H}$, for instance, will be of the form

$$E(\boldsymbol{n}) = \sum_{\alpha} \varepsilon_{\alpha} n_{\alpha} \tag{4.5}$$

where $n_{\alpha} = 0, 1$ for Fermions or $n_{\alpha} = 0, 1, 2, \ldots$ for Bosons.

4.2 The tight-binding model revisited

Let us now apply this to the tight-binding model (3.46):

$$\mathcal{H} = -g \sum_{i=1}^{N} \left(a_i^\dagger a_{i+1} + a_{i+1}^\dagger a_i \right).$$

The matrix H in this case is this case is the same as we studied in Sec. 2.2; e.g., Eq. (2.18):

$$H = -g \begin{pmatrix} 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

To diagonalize $\mathcal{H}$ we therefore first need to import all we learned from Sec. 2.2 about the single-particle tight-binding model. As we have seen, H is diagonalized by introducing a Fourier transform:

$$U_{ik} = \frac{e^{ikx_i}}{\sqrt{N}}, \tag{4.6}$$

where $x_i = i$ with $i = 1, 2, 3, \ldots$ (I set the lattice spacing to 1 and write x_i just so that we don't confuse with the imaginary unit). Moreover the values of k are quantized as in Eq. (2.27):

$$k = \frac{2\pi\ell}{N}, \qquad -\frac{N}{2} < \ell \leq \frac{N}{2}. \tag{4.7}$$

The numbers k are going to play the role here of the eigenvalue index α. In fact, the energy eigenvalues, as we have seen, are

$$\varepsilon_k = -2g \cos k. \tag{4.8}$$

I will also be careful not to mix i, j with k, q. The former refer to positions, with $i, j = 1, 2, \ldots, N$; the latter refer to momenta, with k, q quantized as in (4.7).

Before proceeding, let us verify that the matrix U in Eq. (4.6) is indeed unitary:

$$(U^\dagger U)_{k,q} = \sum_j U_{jk}^* U_{jq} = \frac{1}{N} \sum_{j=1}^{N} e^{i(k-q)x_j}.$$

The resulting sum is nothing but a finite geometric series:

$$\sum_{j=1}^{N} e^{i(k-q)x_j} = \sum_{j=1}^{N} e^{i\theta j} = e^{i\theta} \left(\frac{e^{i\theta N} - 1}{e^{i\theta} - 1} \right), \tag{4.9}$$

where $\theta = k - q$. But because of Eq. (4.7), the term in the numerator involves $e^{i\theta N} = e^{i2\pi(\ell - \ell')} = 1$. The sum therefore vanishes identically. The only exception is when $k = q$, in which case we get $e^{i(k-q)x_j} = 1$ so that the sum reduces to adding N times the number 1. We thus conclude that, indeed

$$(U^\dagger U)_{k,q} = \delta_{kq}, \qquad \to \qquad U^\dagger U = 1,$$

so that U is indeed unitary. I will leave it for you to show that $UU^\dagger = 1$ as well. This also follows from a finite geometric series like (4.9); it is only a bit more annoying because the values of k don't range from $1, \ldots, N$. But you can translate them to do so.

Returning now to our recipe, now that we have the diagonal structure (4.2) of H, the next step is to introduce a new set of creation and annihilation operators as in Eq. (4.3):

$$b_k = \frac{1}{\sqrt{N}} \sum_i e^{ikx_i} a_i, \qquad a_i = \frac{1}{\sqrt{N}} \sum_k e^{-ikx_k} b_k. \tag{4.10}$$

The tight-binding Hamiltonian then becomes

$$\mathcal{H} = \sum_k \varepsilon_k b_k^\dagger b_k, \tag{4.11}$$

which is in diagonal form.

Dealing with k sums:

Once we have a Hamiltonian in the form (4.11), we can proceed to compute physical observables, such as the ground-state, the average energy and so on. To do so we must deal with sums over the discretized k-values (4.7). If the number of sites is small, these sums are difficult to handle. But if N becomes large, we can approximate the sums to an integral. There is a neat trick to do this. Consider a generic sum of the form

$$\sum_k f(k) \tag{4.12}$$

where $f(k)$ is an arbitrary function. Due to the discretization (4.7), the values of k are spaced by

$$\Delta k = \frac{2\pi}{N}, \tag{4.13}$$

which become infinitesimally small when N becomes large. We then multiply Eq. (4.12) by the **convenient 1**: namely $1 = \frac{N}{2\pi}\Delta k$. This leads to

$$\sum_k f(k) = \frac{N}{2\pi} \sum_k \Delta k f(k).$$

When $N \to \infty$, or $\Delta k \to 0$, this approaches exactly the definition of a **Riemann sum** that we learn in Introductory Calculus. Whence, we find

$$\boxed{\sum_k f(k) = \frac{N}{2\pi} \int \mathrm{d}k \; f(k)} \tag{4.14}$$

This is the recipe for converting sums to integrals: always multiply by the "convenient 1". The limits of integration will be $[-\pi, \pi]$ in the case of the tight-binding model, or $-\infty$ to ∞ in the case of free particles.

If we wish to consider spin 1/2 fermions, the result is identical. All we need to do is change the original Hamiltonian to

$$\mathcal{H} = -g \sum_{i,\sigma} \left(a_{i,\sigma}^\dagger a_{i+1,\sigma} + a_{i+1,\sigma}^\dagger a_{i,\sigma} \right),$$

which now contains also a sum over $\sigma = \pm 1$. But since the hopping part does not depend on σ at all, we just need to carry the index σ around everywhere. This then leads to the diagonal Hamiltonian

$$\mathcal{H} = \sum_{k,\sigma} \varepsilon_k b^\dagger_{k,\sigma} b_{k,\sigma}. \tag{4.15}$$

Ground-state in the Fermionic case

As an application of this formula, let us find the Fermi momentum (the momentum of the highest filled state). To this end we need to distinguish between N, which is the number of sites, and N_p, which is the number of particles in the system. We can have, for instance, $N = 10$ sites filled with $N_p = 3$ particles. Let us assume for concreteness that the Fermions have spin 1/2, so that the Hamiltonian is given by Eq. (4.15). In this case each site can take on at most 2 Fermions, so that $N_p \leq 2N$. The Fermi momentum is defined implicitly by

$$\sum_{|k|<k_F,\sigma} = 2 \sum_k \theta(k_F - |k|) = N_p, \tag{4.16}$$

where $\theta(x)$ is the Heaviside function: $\theta(x) = 1$ if $x > 0$ and $\theta(x) = 0$ if $x < 0$. In words, this formula means means that k_F is the value of k which is such that, if you sum up to k_F you get N_p particles. Converting this to an integral we find

$$2\frac{N}{2\pi} \int_{-\pi}^{\pi} \mathrm{d}k \theta(k_F - |k|) = N_p.$$

We can also write this as

$$\frac{N}{2\pi} 4 \int_0^{k_F} \mathrm{d}k = \frac{N}{\pi} k_F = N_p,$$

which leads to

$$k_F = \pi N_p / N := \frac{\pi}{2} n_p. \tag{4.17}$$

The Fermi momentum depends on the **particle density** $n_p \in [0, 2]$.

The dispersion relation (4.8) filled with different values of n_p is illustrated in Fig. 12. It is customary to use as reference the case of **half-filling**, where $N_p = N$ so $n_p = 1$. When the system is at half-filling, we fill exactly half of the band. The Fermi energy is

$$\varepsilon_F = -2g \cos k_F = -2g \cos(\pi n_p / 2), \tag{4.18}$$

which therefore depends non-linearly on n_p.

Let us also compute the ground-state energy. It is defined as

$$E_{\mathrm{gs}} = \sum_{|k|<k_F,\sigma} \varepsilon_k \tag{4.19}$$

Transforming the sum to an integral and using the fact that $\varepsilon_k = -2g \cos k$ is an even function of k, we get

$$E_{\mathrm{gs}} = -g \frac{4N}{\pi} \int_0^{k_F} \mathrm{d}k \cos k.$$

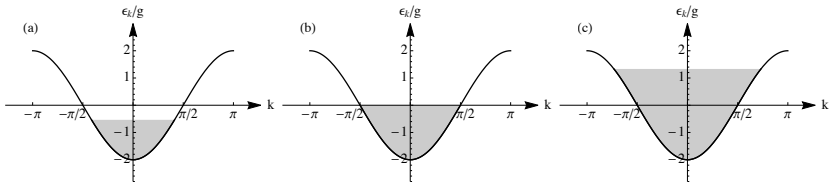

Figure 12: Examples of the energy level fillings for the tight-binding model with different particle densities. The curves represent the dispersion relation (4.8) for (a) below, (b) at and (c) above half-filling $n_p = 1$.

There are a bunch of "2"s in this formula: there is a 2 from the spin degeneracy and a 2 from changing the integral to go from 0 to k_F. There is also a 2 from $-2g\cos k$ and a 2 in the denominator of Eq. (4.14). Carrying out the integration we find

$$E_{\rm gs} = -\frac{4Ng}{\pi}\sin k_F. \tag{4.20}$$

Since $k_F \in [0, \pi/2]$ [Eq. (4.17)], this ranges from 0 all the way to $-4Ng/\pi$ at half-filling $n_p = 1$ and then back to zero when $n_p = 2$.

4.3 Electrons and holes

Now I want to show you something pretty neat. We continue to consider the same Fermionic tight-binding chain as before, with particle density n_p. We then introduce the following changes in nomenclature. For all operators $b_{k\sigma}$ with $|k| \le k_F$ we define

$$c_{k\sigma} = b^\dagger_{k\sigma}, \qquad |k| \le k_F. \tag{4.21}$$

Weird. I know. But let's see what happens. The idea is that we now focus on two sets of operators: $b_{k\sigma}$, which are defined only for $|k| > k_F$ and $c_{k\sigma}$, which are defined for $|k| \le k_F$. If the $b_{k\sigma}$ represent electrons in a lattice, then $b^\dagger_{k\sigma}$ is interpreted as the operator which creates an electron with momentum k. For $|k| \le k_F$, however, we don't say that $b^\dagger_{k\sigma}$ creates an electron. Instead, we say that $c_{k\sigma}$ **annihilates a hole**.

The hole is the electron's "anti-particle". Creating a hole is the same as annihilating an electron and annihilating a hole is the same as creating an electron. Everything below k_F is filled with electrons. We call this the **Fermi sea**. When we apply $c^\dagger_{k\sigma}$ we are removing an electron from the Fermi sea. But instead we interpret this as saying that we are actually creating a hole. Isn't this awesome? :)

Let's look at the algebra. I want to convince that the algebra remains fine. Between the b operators nothing changes:

$$\{b_{k\sigma}, b_{k'\sigma'}\} = 0, \qquad \{b_{k\sigma}, b^\dagger_{k'\sigma'}\} = \delta_{kk'}\delta_{\sigma\sigma'}.$$

For the $c_{k\sigma}$ operators, because Eq. (4.21) just changes the meaning of creation and annihilation, the algebra also remains intact:

$$\{c_{k\sigma}, c_{k'\sigma'}\} = 0, \qquad \{c^\dagger_{k\sigma}, c_{k'\sigma'}\} = \delta_{kk'}\delta_{\sigma\sigma'}.$$

Finally, if we now consider the cross algebra between c's and b's, they will always anti-commute because one set is defined only for $|k| \le k_F$ and the other only for $|k| > k_F$:

$$\{c_{k\sigma}, b_{k\sigma}\} = 0, \qquad \{c_{k\sigma}, b^\dagger_{k\sigma}\} = 0.$$

Next let us look at the number operator:

$$\hat{\mathcal{N}} = \sum_{k\sigma} b^\dagger_{k\sigma} b_{k\sigma} = \sum_{|k|\leq k_F,\sigma} b^\dagger_{k\sigma} b_{k\sigma} + \sum_{|k|>k_F,\sigma} b^\dagger_{k\sigma} b_{k\sigma}.$$

In the first term we use Eq. (4.21) to write

$$\sum_{|k|\leq k_F,\sigma} b^\dagger_{k\sigma} b_{k\sigma} = \sum_{|k|\leq k_F,\sigma} c_{k\sigma} c^\dagger_{k\sigma} = \sum_{|k|\leq k_F,\sigma} (1 - c^\dagger_{k\sigma} c_{k\sigma}),$$

where I used the fermionic algebra to write $c_{k\sigma} c^\dagger_{k\sigma} = 1 - c^\dagger_{k\sigma} c_{k\sigma}$. The first term in the sum is exactly the definition of the Fermi momentum in Eq. (4.16). Hence, the number operator may be written as

$$\boxed{\hat{\mathcal{N}} = N_p + \sum_{|k|>k_F,\sigma} b^\dagger_{k\sigma} b_{k\sigma} - \sum_{|k|\leq k_F,\sigma} c^\dagger_{k\sigma} c_{k\sigma} = N_p + \hat{\mathcal{N}}_e - \hat{\mathcal{N}}_h.} \tag{4.22}$$

The number operator is centered around N_p (the actual number of particles). Electrons count positively to $\hat{\mathcal{N}}$, whereas holes count negatively. Holes are indeed anti-particles! We also do the same for the Hamiltonian (4.15). I will leave for you as an exercise to show that it can be written as

$$\mathcal{H} = E_{\text{gs}} + \sum_{|k|>k_F,\sigma} \varepsilon_k b^\dagger_{k\sigma} b_{k\sigma} - \sum_{|k|\leq k_F,\sigma} \varepsilon_k c^\dagger_{k\sigma} c_{k\sigma}. \tag{4.23}$$

Using this Fermi sea idea, we can really picture the lattice as being populated by **two species**, electrons and holes. In the tight-binding model these two species do not interact with each other. This is visible in Eq. (4.23), where the Hamiltonian is just the sum of the Hamiltonians of the two species. But if we add additional ingredients to the model, the two species will begin to interact.

5 Field quantization

5.1 The Schrödinger field

The operator $a^\dagger_\alpha$ creates a particle at the single-particle state $|\alpha\rangle$. Well, position is also a single-particle state $|x\rangle$. Then why not define an operator $a^\dagger_x$ which creates a particle at position $|x\rangle$? Yeah, we can definitely do that. Except that we don't call it a_x; we use a cooler symbol:

$$\boxed{a_x := \psi(x) = \text{annihilation operator for a particle at position x.}} \tag{5.1}$$

There is a neat reason as to why we use the sacred symbol ψ of wavefunctions. As we will see, $\psi(x)$ does behave a lot like wavefunctions. Except that it is now an operator. This is why we call this **second** quantization: it is a little bit like we are "quantizing" the wavefunction itself. In first quantization we quantize position and momentum. In second quantization, we also quantize the wavefunction. This is not very precise and the name stuck mostly for historical reasons. But that is kind of the logic.

If the system has spin, we upgrade ψ with an additional *internal index* $\psi_\sigma(x)$. This can be $\sigma = \pm 1$ in the case of spin 1/2 or it can be an arbitrary spin value. The commutation relations of the ψ_σ are obtained using the general result (3.22). Since position kets satisfy $\langle x|x'\rangle = \delta(x - x')$ we get

$$[\psi_\sigma(x), \psi^\dagger_{\sigma'}(x')] = \delta_{\sigma,\sigma'}\delta(x - x'), \qquad \text{(Bosons)} \tag{5.2}$$

$$\{\psi_\sigma(x), \psi^\dagger_{\sigma'}(x')\} = \delta_{\sigma,\sigma'}\delta(x - x'). \qquad \text{(Fermions)} \tag{5.3}$$

What about Hamiltonians? Consider first the case of non-interacting systems. The typical single-particle Hamiltonian has the form

$$H_1 = \frac{p^2}{2m} + V(x), \tag{5.4}$$

where $V(x)$ is some external potential. As we learn in undergraduate quantum mechanics, if we move to the position representation we get

$$\langle x|H_1|\phi\rangle = \Big[-\frac{1}{2m}\partial_x^2 + V(x)\Big]\langle x|\phi\rangle,$$

for any wavefunction $|\phi\rangle$. Choosing $|\phi\rangle = |x'\rangle$ we then get

$$\begin{aligned}\langle x|H_1|x'\rangle &= -\frac{1}{2m}\frac{\partial^2}{\partial x^2}\delta(x - x') + V(x)\delta(x - x') \\ &= -\frac{1}{2m}\frac{\partial^2}{\partial x'^2}\delta(x - x') + V(x)\delta(x - x'),\end{aligned} \tag{5.5}$$

where, in the second line, the only thing I did was change the derivative from x to x'. This is allowed because $\frac{\partial}{\partial x}\delta(x - x') = -\frac{\partial}{\partial x'}\delta(x - x')$. But when we differentiate twice, the minus sign goes away.

The general second-quantized version of the single-particle Hamiltonian (5.4) can then be readily found from the recipe in Eq. (3.30):

$$\begin{aligned}\mathcal{H} &= \int \mathrm{d}x\,\mathrm{d}x'\langle x|H_1|x'\rangle\psi^\dagger(x)\psi(x') \\ &= \int \mathrm{d}x\,\mathrm{d}x'\Big[-\frac{1}{2m}\frac{\partial^2}{\partial x'^2}\delta(x - x') + V(x)\delta(x - x')\Big]\psi^\dagger(x)\psi(x').\end{aligned}$$

The term involving $\partial_x^2\delta(x - x')$ is somewhat awkward. But we can get rid of it by integrating by parts. In this case integrating by parts means moving a derivative from one side to another. There are no cross terms because these are evaluated at the boundaries of the box and we use periodic boundary conditions, so any boundary terms vanish. Integrating by parts once:

$$\int \mathrm{d}x\,\mathrm{d}x'\Big[\frac{\partial^2}{\partial x'^2}\delta(x - x')\Big]\psi^\dagger(x)\psi(x') = -\int \mathrm{d}x\,\mathrm{d}x'\Big[\frac{\partial}{\partial x'}\delta(x - x')\Big]\Big[\psi^\dagger(x)\frac{\partial}{\partial x'}\psi(x')\Big]$$

This is the logic of integration by parts. I know this is not very rigorous, But one can arrive at the same result in a more formal way. I promise! Integrating by parts again we get

$$\int \mathrm{d}x\,\mathrm{d}x'\Big[\frac{\partial^2}{\partial x'^2}\delta(x - x')\Big]\psi^\dagger(x)\psi(x') = \int \mathrm{d}x\,\mathrm{d}x'\delta(x - x')\Big[\psi^\dagger(x)\frac{\partial^2}{\partial x'^2}\psi(x')\Big].$$

Now that the delta function is free, we can use its property to eliminate the integral in x or x'.

As a result, we then finally obtain

$$\mathcal{H} = \int \mathrm{d}x\, \psi^\dagger(x)\Big[-\frac{1}{2m}\frac{\partial^2}{\partial x^2} + V(x)\Big]\psi(x). \tag{5.6}$$

This is the second quantized version of a non-interacting Hamiltonian. It is exactly the same as the general recipe (3.30), but specialized to the case of position eigenkets. This introduces the peculiarity that it contains only one integral (which plays the role of the sum), whereas (3.30) contains two.

We can also include interactions in the same spirit. As we saw before, interactions involve more than two operators. For instance, the analog of the Bose-Hubbard Hamiltonian (3.50) is

$$\mathcal{H} = \int \mathrm{d}x\, \psi^\dagger(x)\Big[-\frac{1}{2m}\frac{\partial^2}{\partial x^2}+V(x)\Big]\psi(x)+\frac{U}{2}\int \mathrm{d}x\, \psi^\dagger(x)\psi^\dagger(x)\psi(x)\psi(x). \tag{5.7}$$

This is the Hamiltonian modeling a **superfluid**. The term $V(x)$ refers to an external potential where the particles are trapped, whereas the last term is their Coulomb repulsion.

The free particle revisited

Consider the free particle in a box, obtained by setting $V(x) = 0$ in Eq. (5.6). The second quantized Hamiltonian is then

$$\mathcal{H} = \int \mathrm{d}x\, \psi^\dagger(x)\Big[-\frac{1}{2m}\frac{\partial^2}{\partial x^2}\Big]\psi(x). \tag{5.8}$$

We now introduce a Fourier transform

$$\psi(x)^\dagger = \frac{1}{\sqrt{L}}\sum_k e^{ikx}a_k^\dagger, \qquad k = \frac{2\pi\ell}{L}, \qquad \ell = 0, \pm1, \pm2, \ldots. \tag{5.9}$$

Unlike in tight-binding, here the momentum can take on an infinite number of values ranging from $-\infty$ to ∞.

Introducing (5.9) in (5.8) we get the following:

$$\frac{\partial^2}{\partial x^2}\psi(x) = \frac{1}{\sqrt{L}}\sum_k (-ik)^2 e^{-ikx}a_k$$

Thus,

$$\mathcal{H} = \frac{1}{L}\int \mathrm{d}x \sum_{k,k'} e^{i(k'-k)x}(-ik)^2 a_{k'}^\dagger a_k.$$

Integrating over x and using

$$\frac{1}{L}\int \mathrm{d}x\, e^{i(k'-k)x} = \delta_{k,k'}, \tag{5.10}$$

we finally get

$$\mathcal{H} = \sum_k \frac{k^2}{2m}a_k^\dagger a_k, \tag{5.11}$$

which is the free particle result we analyzed before.

5.2 The Schrödinger Lagrangian

It is possible to cast Schrödinger's equation as a consequence of the **principle of least action**, similar to what we do in classical mechanics. This is fun because it formulates quantum mechanics as a classical theory, as weird as that may sound. This allows us to connect second quantization with field theory.

The principle of least action

Let us start with a brief review of classical mechanics. Consider a system described by a set of generalized coordinates q_i and characterized by a Lagrangian $L(q_i, \partial_t q_i)$. The action is defined as

$$S = \int_{t_1}^{t_2} L(q_i, \partial_t q_i)\, \mathrm{d}t. \tag{5.12}$$

The motion of the system is then generated by the principle of least action; ie, by requiring that the actual path should be an extremum of S. We can find the equations of motion (the Euler-Lagrange equations) by performing a tiny variation in S and requiring that $\delta S = 0$ (which is the condition on any extremum point; maximum or minimum). To do that we write $q_i \to q_i + \eta_i$, where $\eta_i(t)$ is supposed to be an infinitesimal distortion of the original trajectory. We then compute

$$\begin{aligned}
\delta S &= S[q_i(t) + \eta_i(t)] - S[q_i(t)] \\
&= \int_{t_1}^{t_2} \mathrm{d}t \sum_i \left\{ \frac{\partial L}{\partial q_i}\eta_i + \frac{\partial L}{\partial(\partial_t q_i)}\partial_t \eta_i \right\} \\
&= \int_{t_1}^{t_2} \mathrm{d}t \sum_i \left\{ \frac{\partial L}{\partial q_i} - \partial_t\left(\frac{\partial L}{\partial(\partial_t q_i)}\right) \right\}\eta_i.
\end{aligned}$$

where, in the last line, I integrated by parts the second term. Setting each term proportional to η_i to zero then gives us the **Euler-Lagrange equations**

$$\boxed{\frac{\partial L}{\partial q_i} - \partial_t\left(\frac{\partial L}{\partial(\partial_t q_i)}\right) = 0.} \tag{5.13}$$

The example you are probably mostly familiar with is the case when

$$L = \frac{1}{2}m(\partial_t q)^2 - V(q), \tag{5.14}$$

with $V(q)$ being some potential. In this case Eq. (5.13) gives Newton's law

$$m\partial_t^2 q = -\frac{\partial V}{\partial q}. \tag{5.15}$$

Another example, which you may not have seen before, but which will be interesting for us, is the case when we write L with both the position q and the momenta p as generalized coordinates; , ie $L(q, \partial_t q, p, \partial_t p)$. For instance,

$$L = p\partial_t q - H(q, p), \tag{5.16}$$

where H is the Hamiltonian function. In this case there will be two Euler-Lagrange equations for the coordinates q and p:

$$\frac{\partial L}{\partial q} - \partial_t\left(\frac{\partial L}{\partial(\partial_t q)}\right) = -\frac{\partial H}{\partial q} - \partial_t p = 0$$
$$\frac{\partial L}{\partial p} - \partial_t\left(\frac{\partial L}{\partial(\partial_t p)}\right) = \partial_t q - \frac{\partial H}{\partial p} = 0.$$

Rearranging, this gives us Hamilton's equations

$$\partial_t p = -\frac{\partial H}{\partial q}, \qquad \partial_t q = \frac{\partial H}{\partial p}. \tag{5.17}$$

Another thing we will need is the **conjugated momentum** π_i associated to a generalized coordinate q_i. It is always defined as

$$\pi_i = \frac{\partial L}{\partial(\partial_t q_i)}. \tag{5.18}$$

For the Lagrangian (5.14) we get $\pi = m\partial_t q$. For the Lagrangian (5.16) we have two variables, $q_1 = q$ and $q_2 = p$. The corresponding conjugated momenta are $\pi(q) = p$ and $\pi(p) = 0$ (there is no momentum associated with the momentum!). Once we have the momentum we may construct the Hamiltonian from the Lagrangian using the Legendre transform:

$$H = \sum_i \pi_i \partial_t q_i - L \tag{5.19}$$

For the Lagrangian (5.14) we get

$$H = \frac{p^2}{2m} + V(q),$$

whereas for the Lagrangian (5.16) we get

$$H = \pi(q)\partial_t q + \pi(p)\partial_t p - L = p\partial_t q + 0 - p\partial_t q + H = H,$$

as of course expected.

When we go from Lagrangian to Hamiltonian, in Eq. (5.19), we can also quantize our theory. In terms of Lagrangians, q_i are classical variables and the Lagrangian may depend on q_i and $\partial_t q_i$. When we go to a Hamiltonian formulation, we must express the Hamiltonian in terms of the coordinates q_k and the associated conjugated momenta π_i. We then promote q_i and π_i to operators satisfying the canonical algebra

$$[q_i, \pi_j] = i\delta_{ij} \tag{5.20}$$

This is the idea of **canonical quantization**.

A principle of least action for Schrödinger's equation

Now consider Schrödinger's equation in first quantization

$$i\partial_t \psi = H\psi, \tag{5.21}$$

where ψ is just the usual c-number wavefunction. We can write this in terms of an arbitrary basis $|n\rangle$ by defining $\psi_n = \langle n|\psi\rangle$. Schrödinger's equation then becomes

$$i\partial_t \psi_n = \sum_m H_{n,m}\psi_m, \tag{5.22}$$

where $H_{n,m} = \langle n|H|m\rangle$. We now ask the following question: can we cook up a Lagrangian and an action such that the corresponding Euler-Lagrange equations give Eq. (5.22)? The answer, of course, is yes.[4] The "variables" in this case are all components ψ_n. But since they are complex variables, we actually have ψ_n and ψ_n^* as an independent set. For reasons which will become clear in a second, I will write $\psi_n^\dagger$ instead of ψ_n^*. At this level this is the same thing since ψ_n is just a c-number. The Lagrangian is then $L = L(\psi_n, \partial_t\psi_n, \psi_n^\dagger, \partial_t\psi_n^\dagger)$. and the action is

$$S[\psi_n^\dagger, \psi_n] = \int_{t_1}^{t_2} L(\psi_n, \partial_t\psi_n, \psi_n^\dagger, \partial_t\psi_n^\dagger)\,\mathrm{d}t. \tag{5.23}$$

The correct Lagrangian we should use is

$$\boxed{L = \sum_n i\psi_n^\dagger \partial_t \psi_n - \sum_{n,m} H_{n,m}\psi_n^\dagger\psi_m.} \tag{5.24}$$

where ψ_n and $\psi_n^\dagger$ are to be interpreted as independent variables. Please take notice of the similarity with Eq. (5.16): ψ_n plays the role of q and $\psi_n^\dagger$ plays the role of p. To check that this works we use the Euler-Lagrange equations for the variable $\psi_n^\dagger$:

$$\frac{\partial L}{\partial \psi_n^\dagger} - \partial_t\left(\frac{\partial L}{\partial(\partial_t\psi_n^\dagger)}\right) = 0.$$

The second term is zero since $\partial_t\psi_n^\dagger$ does not appear in Eq. (5.24). The first term then gives

$$\frac{\partial L}{\partial \psi_n^\dagger} = i\partial_t\psi_n - \sum_m H_{n,m}\psi_m = 0.$$

which is precisely Eq. (5.22). Thus, we have just cast Schrödinger's equation as a principle of least action for a weird action that depends on the quantum state $|\psi\rangle$. I will leave to you as an exercise to compute the Euler-Lagrange equation for ψ_n; you will simply find the complex conjugate of Eq. (5.22).

Eq. (5.24) is written in terms of the components ψ_n of a certain basis. We can also write it in a basis independent way, as

$$L = \langle\psi|(i\partial_t - H)|\psi\rangle \tag{5.25}$$

This is what I call the Schrödinger Lagrangian. Isn't it beautiful?

We can also ask what is the conjugated momentum associated with the variable ψ_n for the Lagrangian (5.24). Using Eq. (5.18) we get,

$$\pi(\psi_n) = \frac{\partial L}{\partial(\partial_t\psi_n)} = i\psi_n^\dagger, \qquad \pi(\psi_n^\dagger) = 0 \tag{5.26}$$

[4] If the answer was no, I would be a completely crazy person, because I just spent more than two pages describing Lagrangian mechanics, which would have all been for nothing.

This means that ψ_n and $i\psi_n^\dagger$ are conjugated variables. As a sanity check, we can now find the Hamiltonian using the definition (5.19):

$$H = \sum_n i\psi_n^\dagger \partial_t \psi_n - L = \sum_{n,m} H_{n,m} \psi_n^\dagger \psi_n. \tag{5.27}$$

which is just the actual Hamiltonian.

Finally, we can write it in terms of the coordinate representation. In this case the action can be expressed in terms of a *Lagrangian density* as

$$S[\psi^*, \psi] = \int \mathrm{d}t\, \mathrm{d}x\, \mathcal{L}(\psi, \partial_t \psi, \psi^*, \partial_t \psi^*), \tag{5.28}$$

where

$$\boxed{\mathcal{L} = i\psi^\dagger(x)\partial_t \psi(x) - \psi^\dagger(x)\Big[-\frac{1}{2m}\frac{\partial^2}{\partial x^2} + V(x)\Big]\psi(x).} \tag{5.29}$$

Written in this way, Schrödinger equation is thus seen to be a **classical field theory** for the field $\psi(x)$.

We can now **quantize the Schrödinger field** using the canonical quantization procedure. I know this is a bit weird, but let's see what comes out of it. The procedure is to (i) move to the Hamiltonian representation, (ii) express it in terms of coordinates and conjugated momenta and (iii) promote them to operators satisfying the canonical algebra (5.20). But the momentum π conjugated to $\psi(x)$ and $\psi^\dagger(x)$ are, according to Eq. (5.26), given by $\pi(\psi(x)) = i\psi^\dagger(x)$ and $\pi(\psi^\dagger(x)) = 0$. The former, when combined with the canonical quantization condition (5.20), implies that

$$[\psi(x), i\psi^\dagger(x)] = i\delta(x - x'), \tag{5.30}$$

where I replaced the Kronecker with a Dirac delta. Notice how the factors of i cancel out, leaving us with

$$[\psi(x), \psi^\dagger(x)] = \delta(x - x'), \tag{5.31}$$

which is *exactly* the commutation relations for Bosons in Eq. (5.2). This is a neat way of showing why second quantization is indeed a **second** quantization: we interpret Schrödinger's equation as a classical equation and then requantize it. This promotes the wavefunction to an operator satisfying the canonical algebra (5.31). We could do something analogous to Fermions, by assuming that canonical quantization should involve anti-commutators. Of course, at this level, this has to be done pretty much by hand. The true origin of anti-commutation relations for Fermions can only be justified in quantum field theory.

A More on the Bosonic and Fermionic algebras

A.1 Commutators and anti-commutators

The following formulas are useful to know:

$$[A, BC] = B[A, C] + [A, B]C \tag{A.1}$$

$$[AB, C] = A[B, C] + [A, C]B \tag{A.2}$$

There is an easy way to remember them. For instance, in $[A, BC]$ you first take B out to the left and then C out to the right.

For Fermionic operators we sometimes like to convert commutators into anti-commutators:

$$[A, BC] = \{A, B\}C - B\{A, C\} \tag{A.3}$$

$$[AB, C] = A\{B, C\} - \{A, C\}B \tag{A.4}$$

You can remember this as follows: the logic is the same as in Eqs. (A.1) and (A.2), but for anti-commutators, whenever an operator need to jump over someone to escape, it gains a minus sign. For 4 operators we have

$$[AB, CD] = -AC\{D, B\} + A\{C, B\}D - C\{D, A\}B + \{C, A\}DB \tag{A.5}$$

A.2 Eigenstuff of bosonic operators

In this appendix I summarize a calculation found in any quantum mechanics book: *Given a non-Hermitian operator a satisfying*

$$[a, a^\dagger] = 1, \tag{A.6}$$

what are the eigenvalues and eigenvectors of $a^\dagger a$. This is a really important problem that appears often in all areas of quantum physics: given an algebra, find the eigenstuff. I think this is one of those things that everyone should go through once. So here we go.

Since $a^\dagger a$ is Hermitian, its eigenvalues must be real and its eigenvectors can be chosen to form an orthonormal basis. Let us write them as

$$a^\dagger a|n\rangle = n|n\rangle. \tag{A.7}$$

Our goal is to find the allowed n and the corresponding $|n\rangle$. We of course already know that n is an integer. But let's pretend we did not know and let's figure out how this follows from the algebra.

The first thing we notice is that $a^\dagger a$ must be positive semi-definite operator, so n cannot be negative:

$$n = \langle n|a^\dagger a|n\rangle \geq 0.$$

Next we use Eq. (A.6) to show that

$$[a^\dagger a, a] = -a, \qquad [a^\dagger a, a^\dagger] = a^\dagger. \tag{A.8}$$

This type of structure is a signature of a *ladder*-like spectrum (that is, when the eigenvalues are equally spaced). To see that, we use these commutation relations to compute:

$$(a^\dagger a)a|n\rangle = [a(a^\dagger a) - a]|n\rangle = a(a^\dagger a - 1)|n\rangle = (n-1)a|n\rangle. \tag{A.9}$$

Hence, we conclude that if $|n\rangle$ is an eigenvector with eigenvalue n, then $a|n\rangle$ is also an eigenvector, but with eigenvalue $(n-1)$ [This is the key argument. Make sure you understand what this sentence means.]. However, I wouldn't call this $|n-1\rangle$ just yet because $a|n\rangle$ is not normalized. Thus we need to write

$$|n-1\rangle = \gamma a|n\rangle,$$

where γ is a normalization constant. To find it we simply write

$$1 = \langle n-1|n-1\rangle = |\gamma|^2\langle n|a^\dagger a|n\rangle = |\gamma|^2 n.$$

Thus $|\gamma|^2| = 1/n$. The phase of γ is arbitrary so we choose it for simplicity as being real and positive. We then get

$$|n-1\rangle = \frac{a}{\sqrt{n}}|n\rangle.$$

From this analysis we conclude that a reduces the eigenvalues by unity:

$$a|n\rangle = \sqrt{n}|n-1\rangle.$$

We can do a similar analysis with $a^\dagger$. We again use Eq. (A.8) to compute

$$(a^\dagger a)a^\dagger|n\rangle = (n+1)a^\dagger|n\rangle.$$

Thus $a^\dagger$ raises the eigenvalue by unity. The normalization factor is found by a similar procedure: we write $|n+1\rangle = \beta a^\dagger|n\rangle$, for some constant β, and then compute

$$1 = \langle n+1|n+1\rangle = |\beta|^2\langle n|aa^\dagger|n\rangle = |\beta|^2\langle n|(1+a^\dagger a)|n\rangle = |\beta|^2(n+1).$$

Thus

$$a^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle.$$

Now comes the trickiest (and most beautiful) argument. We have seen that if n is an eigenvalue, then $n \pm 1$, $n \pm 2$, etc., will all be eigenvalues. But this doesn't mean that n itself should be an integer. Maybe we find one eigenvalue which is 42.3; then 41.3, 43.3 and so on will also be eigenvalues. But suppose we start with some eigenstate $|n\rangle$ and keep on applying a a bunch of times. At each application we will lower the eigenvalue by one tick:

$$a^\ell|n\rangle = \sqrt{n(n-1)\dots(n-\ell+1)}|n-\ell\rangle.$$

But this party cannot continue forever because, as we have just discussed, the eigenvalues of $a^\dagger a$ cannot be negative. They can, at most, be zero. The only way for this to happen is if there exists a certain *integer* ℓ for which $a^\ell|n\rangle \neq 0$ but $a^{\ell+1}|n\rangle = 0$. And this can only happen if $\ell = n$ because, then

$$a^{\ell+1}|n\rangle = \sqrt{n(n-1)\dots(n-\ell+1)(n-\ell)}|n-\ell-1\rangle = 0,$$

and the term $n-\ell$ will vanish. Since ℓ is an integer, we therefore conclude that n must also be an integer. Thus, we finally conclude that

$$\text{eigs}(a^\dagger a) = n \in \{0,1,2,3,\dots\}. \tag{A.10}$$

It is for this reason that $a^\dagger a$ is called the number operator: we usually say $a^\dagger a$ counts the number of quanta in a given state: given a state $|n\rangle$, you first apply a to annihilate one quant and then $a^\dagger$ to create it back again. The proportionality factor is the eigenvalue n. Curiously, this analysis seem to imply that if you want to count how many people there are in a room, you first need to annihilate one person and then create a fresh new human. Quantum mechanics is indeed strange.

This analysis also serves to define the **vacuum** state, with $n = 0$; it satisfies

$$\boxed{a|0\rangle = 0.} \tag{A.11}$$

We can build all states starting from the vacuum and applying $a^\dagger$ successively:

$$|n\rangle = \frac{(a^\dagger)^n}{\sqrt{n!}}|0\rangle. \tag{A.12}$$

Using this and the algebra of a and $a^\dagger$ it then follows that the states $|n\rangle$ form an orthonormal basis, as expected:

$$\langle n|m\rangle = \delta_{n,m}.$$

A.3 Eigenstuff of fermionic operators

Next let us do the same for fermionic operators, which satisfy

$$\{a, a^\dagger\} = 1, \qquad (a^\dagger)^2 = 0. \tag{A.13}$$

Using Eq. (A.4) you may verify that Eqs. (A.8) continue to hold:

$$[a^\dagger a, a] = -a, \qquad [a^\dagger a, a^\dagger] = a^\dagger. \tag{A.14}$$

Now I want you to stop and think for a second. If you go back to Sec. A.2 you will notice that the entire derivation was not based on $[a, a^\dagger] = 1$; it was based on Eq. (A.14). Thus, if this relation also holds for Fermions, everything else must hold identically as well. In particular, there is also a vacuum which is annihilated by a [Eq. (A.11)] and from which all other states can be constructed Eq. (A.12). The *only* difference is that for Fermions $(a^\dagger)^2 = 0$. Whence, we cannot build an infinite number of states from $|0\rangle$, but only one:

$$a^\dagger|1\rangle = (a^\dagger)^2|0\rangle = 0. \tag{A.15}$$

As a consequence, the eigenvectors of $a^\dagger a$ in the case of Bosons is composed solely of two vectors, $|0\rangle$ and $|1\rangle$.