

The second quantization

In this section we discuss why working with wavefunctions is not a good idea in systems with $N \sim 10^{23}$ particles, and introduce much more suitable notation for this purpose.

1 Quantum mechanics for one particle - review of notation

The dynamics of the particle is described by some Hermitian Hamiltonian \hat{H} , acting on the one-particle Hilbert space \mathcal{H} . The Hilbert space is generally a product of Hilbert spaces associated with the translational and the spin degrees of freedom, $\mathcal{H} = \mathcal{H}_{tr} \otimes \mathcal{H}_{spin}$. However, sometimes we will “forget” about the spin degrees of freedom, and be concerned only with the location in space of the so-called “spinless” particle. In this case \hat{H} contains no spin operators. Other times we will “forget” the translational degrees of freedom and concern ourselves only with the dynamics of localized spins – this works, for instance, in insulators with non-trivial magnetic properties.

Examples of one-particle Hamiltonians:

$$\hat{H} = \frac{\hat{p}^2}{2m} \rightarrow \text{free (possibly spinless) particle;}$$

$$\hat{H} = \frac{\hat{p}^2}{2m} + V(\vec{r}, t) \rightarrow \text{particle in an external field.}$$

$$\hat{H} = \frac{1}{2m} \left[\hat{p} - q\vec{A}(\vec{r}, t) \right]^2 + q\Phi(\vec{r}, t) - g\mu_B \hat{\vec{S}} \cdot \vec{B}(\vec{r}, t) \rightarrow \text{spinful particle of charge } q \text{ in an external classical electromagnetic field, } \vec{B} = \nabla \times \vec{A}, \vec{E} = -\nabla\Phi - \frac{\partial \vec{A}}{\partial t}.$$

$$\hat{H} = -\gamma \hat{\vec{S}} \cdot \vec{B}(t) \rightarrow \text{localized particle of spin } S \text{ in external magnetic field } \vec{B}(t).$$

The list goes on and on. **Note:** all quantities with a hat are operators acting on the Hilbert space (sometimes trivially).

We can always work in an abstract Hilbert space, with abstract states denoted by $|\psi\rangle$. Sometimes, however, it is convenient to use particular representations, such as the \vec{r} -representation. In this case, we use the complete, orthonormal basis set for the Hilbert space:

$$|\vec{r}, S_z\rangle = |\vec{r}\rangle \otimes |S, S_z\rangle$$

where $\hat{r}|\vec{r}\rangle = \vec{r}|\vec{r}\rangle$ are eigenstates of the position operator, while $\hat{\vec{S}}^2 |S, S_z\rangle = \hbar^2 S(S+1) |S, S_z\rangle$, $\hat{S}_z |S, S_z\rangle = \hbar S_z |S, S_z\rangle$ are eigenstates of the spin operators $\hat{\vec{S}}^2, \hat{S}_z$. It is convenient to introduce the simplified notation $x = (\vec{r}, S_z)$ such that $|x\rangle = |\vec{r}, S_z\rangle$. In this case, the orthonormalization and the completeness relations are written as:

$$\langle x|x'\rangle = \langle \vec{r}|\vec{r}'\rangle \langle S, S_z|S, S'_z\rangle = \delta(\vec{r} - \vec{r}') \delta_{S_z, S'_z} = \delta_{x, x'} \quad (1)$$

and

$$\mathbf{1} = \sum_{S_z=-S}^S \int d\vec{r} |\vec{r}, S_z\rangle \langle \vec{r}, S_z| = \sum dx |x\rangle \langle x| \quad (2)$$

The last notation should hopefully remind us to sum over all discrete variables, and integrate over all continuous ones. In this basis, any state can be decomposed as:

$$|\psi\rangle = \sum dx |x\rangle \langle x|\psi\rangle = \sum_{S_z=-S}^S \int d\vec{r} \psi(\vec{r}, S_z) |\vec{r}, S_z\rangle$$

where $\psi(x) = \psi(\vec{r}, S_z) = \langle \vec{r}, S_z | \psi \rangle$, also known as the wavefunction, is the amplitude of probability that the particle in state $|\psi\rangle$ is at position \vec{r} with spin S_z . This amplitude is generally a complex number. **Note:** if the particle has a spin $S > 0$, then $\langle \vec{r} | \psi \rangle$ is a spinor with $2S + 1$ entries,

$$\langle \vec{r} | \psi \rangle = \psi(\vec{r}) = \begin{pmatrix} \psi(\vec{r}, S) \\ \psi(\vec{r}, S-1) \\ \vdots \\ \psi(\vec{r}, -S) \end{pmatrix}$$

Then, for instance, the probability to find the particle at \vec{r} is $\psi^\dagger(\vec{r})\psi(\vec{r}) = \sum_{S_z} |\psi(\vec{r}, S_z)|^2$, as expected. Clearly, if we know the wavefunctions $\psi(x) \forall x$, we know the state of the system. Any equation for $|\psi\rangle$ can be turned into an equation for $\psi(x)$ by simply acting on it with $\langle x|$, for example: $|\psi_1\rangle = |\psi_2\rangle \rightarrow \psi_1(x) = \psi_2(x)$. However, we also have to deal with expression of the form $\langle x | \hat{A} | \psi \rangle$, where \hat{A} is some abstract operator. In fact, the only operators that appear in the Hamiltonian are (combinations of):

$$\langle x | \hat{r} | \psi \rangle = \vec{r} \psi(\vec{r}, S_z)$$

$$\langle x | \hat{p} | \psi \rangle = -i\hbar \nabla \psi(\vec{r}, S_z)$$

$$\langle x | \hat{S}_z | \psi \rangle = \hbar S_z \psi(\vec{r}, S_z); \quad \langle x | \hat{S}_\pm | \psi \rangle = \hbar \sqrt{(S \mp S_z)(S \pm S_z + 1)} \psi(\vec{r}, S_z \pm 1)$$

where $\hat{S}_\pm = \hat{S}_x \pm i\hat{S}_y$. In general, $\langle x | \hat{A} | \psi \rangle = \sum dx' \langle x | \hat{A} | x' \rangle \psi(x')$. As long as we can compute the matrix elements of the operator in the basis of interest, we have its representation in that basis.

Again, note that if we project on $\langle \vec{r} |$ only, these are equations for spinors. In this case, the last equation changes to a matrix equation

$$\langle \vec{r} | \hat{S}_z | \psi \rangle = \tilde{S}_z \psi(\vec{r}) = \begin{pmatrix} \hbar S & 0 & \dots & 0 \\ 0 & \hbar(S-1) & \dots & 0 \\ \dots & \dots & \dots & 0 \\ 0 & 0 & 0 & -\hbar S \end{pmatrix} \begin{pmatrix} \psi(\vec{r}, S) \\ \psi(\vec{r}, S-1) \\ \dots \\ \psi(\vec{r}, -S) \end{pmatrix}$$

i.e. each spin-component gets multiplied by its particular spin projection. One can also find the matrix representations for the operators \hat{S}_x, \hat{S}_y (**exercise** - do it!).

Example: consider a spin- $\frac{1}{2}$ particle in an external magnetic field, described by the abstract Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} - \gamma \hat{\vec{S}} \cdot \vec{B}(t)$$

In the \vec{r} -representation, the Schrödinger equation $i\hbar \frac{d}{dt} |\psi(t)\rangle = \hat{H}(t) |\psi(t)\rangle$ can be written in a compact form, as a spinor equation:

$$i\hbar \frac{d\psi(\vec{r})}{dt} = \left[-\frac{\hbar^2}{2m} \nabla^2 - \gamma \hat{\vec{S}} \cdot \vec{B}(t) \right] \psi(\vec{r})$$

where

$$\psi(\vec{r}) = \begin{pmatrix} \psi(\vec{r}, \uparrow) \\ \psi(\vec{r}, \downarrow) \end{pmatrix}$$

and the spin- $\frac{1}{2}$ matrices are $\hat{S}_i = \frac{\hbar}{2} \sigma_i$, where the matrices σ are called the Pauli matrices:

$$\sigma_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}; \quad \sigma_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}; \quad \sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

You should convince yourselves that the two equations for the wavefunctions $\psi(\vec{r}, S_z)$ that we obtain by projecting the abstract equation onto $\langle \vec{r}, S_z |$ are equivalent to this one spinor equation.

All this may seem rather trivial and somewhat of a waste of time. However, it is useful to remember that using wavefunctions $\psi(x)$ is a choice and should be done only when convenient. When not convenient, we may use another representation (for instance of momentum states, i.e. projecting on $|\vec{k}, S_z\rangle$ basis, or coherent states, or any other number of options). Or, we may decide to work directly in the abstract space.

2 Why we need 2nd quantization for systems with $N \gg 1$ particles

As we have seen, we can formulate the quantum mechanics for a single particle in an abstract Hilbert space, although in practice we (almost) always end up working in some given representation, in terms of wavefunctions. Can we generalize this approach to many-body systems? Yes, we can quite easily, if the number N of particles is fixed. In this case, the many-body state in the \vec{r} -representation is (we use again $x = (\vec{r}, \sigma)$):

$$\langle x_1, \dots, x_N | \Psi(t) \rangle = \Psi(x_1, \dots, x_N, t)$$

which is the amplitude of probability that at time t , particle 1 is at $x_1 = (\vec{r}_1, \sigma_1)$, etc. Note that if the number of particles is not fixed (as is the case in a grand-canonical ensemble) we're in trouble deciding what $\langle |$ to use. But let us ignore this for the moment, and see how we get in trouble even if the number of particles is fixed.

We almost always decompose wavefunctions in a given basis. For a single particle, we know that any wavefunction $\Psi(x, t)$ can be decomposed in terms of a complete and orthonormal basis $\phi_\alpha(x) = \langle x | \alpha \rangle$ as $\Psi(x, t) = \sum_\alpha c_\alpha(t) \phi_\alpha(x)$ – this reduces the problem to that of working with the time-dependent complex numbers $c_\alpha(t)$.

If we have a complete basis for a single-particle Hilbert space, we can immediately generate a complete basis for the N -particle Hilbert space (the particles are identical), as the products of one-particle basis states:

$$\Psi(x_1, \dots, x_N, t) = \sum_{\alpha_1, \dots, \alpha_N} c_{\alpha_1, \dots, \alpha_N}(t) \phi_{\alpha_1}(x_1) \phi_{\alpha_2}(x_2) \cdots \phi_{\alpha_N}(x_N)$$

As usual, $c_{\alpha_1, \dots, \alpha_N}(t)$ is the amplitude of probability to find particle 1 in state α_1 and located at x_1 , etc.

However, because the particles are identical we know that the wavefunctions must be symmetric (for bosons) or antisymmetric (for fermions) to interchange of any particles:

$$\Psi(x_1, \dots, x_N, t) = \xi^P \Psi(x_{P_1}, \dots, x_{P_N}, t)$$

where

$$\begin{pmatrix} 1 & 2 & \dots & N \\ P_1 & P_2 & \dots & P_N \end{pmatrix}$$

is any permutation, P is its sign (number of transpositions), and from now on we use the notation:

$$\xi = \begin{cases} -1, & \text{for fermions} \\ +1, & \text{for bosons} \end{cases} \quad (3)$$

It follows immediately that $c_{\dots\alpha_i,\dots,\alpha_j,\dots} = \xi c_{\dots,\alpha_j,\dots,\alpha_i,\dots}$ and so $c_{\dots,\alpha_i,\dots,\alpha_i,\dots} = 0$ for fermions: we cannot have two or more fermions occupying identical states α_i – **the Pauli principle** is automatically enforced through this symmetry.

Because of this requirement, the physically meaningful many-body fermionic (bosonic) wavefunctions are from the antisymmetric (symmetric) sector of the N -particle Hilbert space, and we only need to keep the properly symmetrized basis states, which we denote as:

$$\phi_{\alpha_1,\dots,\alpha_N}(x_1,\dots,x_N) = \sqrt{\frac{1}{N! \prod_i n_i!}} \sum_{P \in S_N} \xi^P \phi_{\alpha_1}(x_{P_1}) \phi_{\alpha_2}(x_{P_2}) \cdots \phi_{\alpha_N}(x_{P_N}) \quad (4)$$

The factor in front is the normalization constant; n_i is the total number of particles in the same state α_i (only important for bosonic systems; in fermionic systems all $n_i = 1$), and the summation is over all possible permutations. You should check that indeed, these functions are properly normalized. For fermions, such a properly antisymmetrized product of one-particle states is called a **Slater determinant**.

The set of all these properly symmetrized/antisymmetrized basis states $\phi_{\alpha_1,\dots,\alpha_N}(x_1,\dots,x_N)$ can be shown to form a complete basis for the N -particle Hilbert space, when the particles are bosons/fermions. Then, any N -body wavefunction can be written as:

$$\Psi(x_1,\dots,x_N,t) = \sum_{\alpha_1,\dots,\alpha_N} c_{\alpha_1,\dots,\alpha_N}(t) \phi_{\alpha_1,\dots,\alpha_N}(x_1,\dots,x_N)$$

This means that even if we are extremely lucky, and only a single combination of one-particle states α_1,\dots,α_N is occupied, so that the sum contains a single $\phi_{\alpha_1,\dots,\alpha_N}(x_1,\dots,x_N)$ basis wavefunction, this alone contains on the order of $N!$ terms (the sum over all permutations in Eq. (4)). If there are more basis states involved in the decomposition of Ψ , then there are that many more terms on the right-hand side of the decomposition.

So now we can see why it is inconvenient to use this approach. Even if the number of particles is fixed (which is usually not the case); and even if we have managed to solve somehow the problem and find the many-body wavefunction $\Psi(x_1,\dots,x_N,t)$ (which is to say, we know its decomposition coefficients $c_{\alpha_1,\dots,\alpha_N}(t)$ for the given basis) – what we really need – for comparison with experiments – are expectation values of single particle operators (such as the total momentum, or particle density, or whatever interests us) or two-particle operators (such as a Coulomb potential interaction). Any single particle operator is of general form $\hat{A} = \sum_{i=1}^N \hat{A}_i$, where \hat{A}_i is the operator acting in the single-particle Hilbert space of particle i . Similarly, two-particle operators are of general form $\hat{B} = \sum_{i<j} \hat{B}_{ij}$, with a total of $N(N-1)/2$ terms in the sum, one for each pair of particles. So what we typically have to calculate is something of the form:

$$\langle \Psi | \mathcal{O} | \Psi \rangle = \sum dx_1 \dots \sum dx_N \Psi^*(x_1,\dots,x_N,t) \mathcal{O} \Psi(x_1,\dots,x_N,t)$$

so that we have to perform N integrals over real space (all operators we deal with in this course are diagonal in positions); in general $2N$ sums over spin indexes (which reduce to N if the operator is diagonal in spin-space, as I assumed here) ... and this is out of a combination of a product of order $(N!)^2$ terms contained in $\Psi\Psi^*$, times N or $N(N-1)/2$ terms from the action of the operator. If $N = 1$ or 2 , which is the case in introductory QM where this approach was taught to us, this is not really a problem. However, if $N \sim 10^{23}$ like we need to deal with for condensed matter systems, it becomes exceedingly cumbersome and unpleasant to keep track of all these many terms. And

as I said, if the number of particles is not fixed, which is the case if we work in a grand-canonical ensemble, then things become that much uglier.

The origin of these complications is the fact that we insisted on working with wavefunctions which contain a lot of useless information (which particle is where). In fact, because the particles are indistinguishable, all we need to know is which one-particle states are occupied, and we do not need to bother listing which particle is where – we know that all possible permutations will appear, anyway. Keeping only this minimal amount of necessary information is precisely what 2nd quantization does.

3 2nd quantization

Notation: from now on, I will use a single index α to label states of a complete single-particle basis (including the spin). Which basis to use depends on the problem at hand: for instance, for translationally invariant problems, we will use $\alpha = (\vec{k}, \sigma)$ as quantum numbers, whereas if we deal with an atom, we can use $\alpha = (n, l, m, \sigma)$ for its one-particle eigenstates (the usual hydrogen-like orbitals). So α is simply a shorthand notation for the collection of all quantum numbers needed to characterize the single-particle state: if there is a single particle in the system, we can identify each of its possible states by a unique set of values for the numbers making up α .

Next, we define an ordering for these states $(\alpha_1, \alpha_2, \dots)$; for instance, we order them in increasing order of energy for some single-particle Hamiltonian $E_{\alpha_1} < E_{\alpha_2} < \dots$ etc. If there are degenerate states, we define some rule to order the states, e.g. spin-up first and spin-down second. You may worry that if \vec{k} is one of the quantum numbers, we cannot order continuous variables – however, we will place the system in a box of volume V so that only discrete \vec{k} values are allowed (which we can order), and then let $V \rightarrow \infty$ at the end of all calculations. So, in practice there is always some way to order these one-particle states.

Once this ordering is agreed upon, we define the abstract vectors:

$$|n_1, n_2, \dots\rangle \quad (5)$$

as being the state with n_1 particles in state “1” with index α_1 , n_2 particles in state “2” with index α_2 , etc. We have to list occupation numbers for all possible states – for an empty state, $n_i = 0$. Of course, for fermionic systems we can only have $n_i = 0$ or 1, for any i (Pauli’s principle). For bosons, $n_i = 0, 1, 2, \dots$ can be any non-negative integer.

Examples (we’ll discuss more of these in class): the ground-state for N non-interacting fermions is represented as $|1, 1, \dots, 1, 0, 0, \dots\rangle$ (first lowest-energy N states are occupied, the other ones are all empty) whereas the ground-state of a non-interacting bosonic system is $|N, 0, 0, \dots\rangle$ (all bosons in the lowest energy state). The vacuum is $|0, 0, \dots\rangle$ in both cases. Etc.

The ensemble of all possible states $\{|n_1, n_2, \dots\rangle\}$ is a complete orthonormal basis of the so-called **Fock space**. The Fock space is the reunion of the Hilbert (fermionic or bosonic) spaces with any number of particles, from zero (the vacuum) to any $N \rightarrow \infty$. If the number of particles is fixed, we work in the Hilbert subspace of the Fock space defined by the constraint $N = \sum_i n_i$.

The link between these abstract states and the N -particle basis states of Eq. (4) is straightforward. For fermions, the Slater determinant $\phi_{\alpha_1, \dots, \alpha_N}(x_1, \dots, x_N) = \langle x_1, \dots, x_N | \dots, n_{\alpha_1} = 1, \dots, n_{\alpha_N} = 1, \dots \rangle$ is the wavefunction associated with the abstract state that has the one-particle states $\alpha_1, \dots, \alpha_N$ occupied, while all other one-particle states are empty; and similarly for bosons, except their occupation numbers can be larger than 1. As advertised, the abstract state contains

only the key information of which one-particle states are occupied. By contrast, the Slater determinants (and their bosonic analogues) also contain the unnecessary information of which particle is in which one-particle state, and where in space it is located.

We would like now to be able to generate easily these abstract states, and also to work with operators represented directly in this Fock space, so that computing their matrix elements is easy.

Remember that these are identical particles, so these basis states must obey the proper statistics. We enforce this in the following way: to each single-particle state α we associate a pair of operators $c_\alpha, c_\alpha^\dagger$ which obey the algebra:

$$\begin{aligned} c_\alpha c_\beta^\dagger - \xi c_\beta^\dagger c_\alpha &= [c_\alpha, c_\beta^\dagger]_\xi = \delta_{\alpha,\beta} \\ [c_\alpha, c_\beta]_\xi &= [c_\alpha^\dagger, c_\beta^\dagger]_\xi = 0 \end{aligned} \quad (6)$$

We use the notation $[a, b]_\xi = ab - \xi ba$ to deal simultaneously with both fermions ($\xi = -1$) and bosons ($\xi = 1$). When dealing with a well-defined type, we will use the customary notation for commutators $[,]_+ = [,]$ and for anticommutators $[,]_- = \{, \}$. Also, in this section I will call fermionic operators as $a_\alpha, a_\alpha^\dagger$, and bosonic operators $b_\alpha, b_\alpha^\dagger$. I will use the c operators when I want to deal with both types simultaneously. These operators are called **creation and annihilation operators**. This is because as we will see next, $a_\alpha^\dagger|\Psi\rangle$ is the state we obtain by adding a particle in state α to the state $|\Psi\rangle$; similarly, $a_\alpha|\Psi\rangle$ is the state we obtain by removing a particle in state α from state $|\Psi\rangle$ (if there is no particle in this state, then $a_\alpha|\Psi\rangle = 0$).

For fermions, Pauli's principle is automatically obeyed because the second line of the Wigner-Pauli algebra (Eq. (6)) gives that: $(a_\alpha^\dagger)^2 = (a_\alpha)^2 = 0$: we can't create or annihilate two fermions in/from the same state.

Let $\hat{n}_\alpha = c_\alpha^\dagger c_\alpha$ - as we will see now, this is **the number operator**, which tells us how many particles are in state α .

For fermions $\hat{n}_\alpha \hat{n}_\alpha = a_\alpha^\dagger a_\alpha a_\alpha^\dagger a_\alpha = a_\alpha^\dagger (1 - a_\alpha^\dagger a_\alpha) a_\alpha = \hat{n}_\alpha$. As a result, the operator \hat{n}_α can only have eigenvalues 0 or 1 (which, of course, is precisely what Pauli said). Let us look at each α individually (and not write its label α , for the moment). Then the eigenstates of this operator are $\hat{n}|0\rangle = 0; \hat{n}|1\rangle = |1\rangle$. Let's now see how the individual operators a, a^\dagger act on these two states. Start from $\hat{n}|0\rangle = a^\dagger a|0\rangle = 0$ and act on both sides with a . Using $aa^\dagger a = (1 - a^\dagger a)a = a$ because $a^2 = 0$, we find $a|0\rangle = 0$. Which hopefully feels re-assuring, as it says that we cannot remove a particle if the state is already empty. Let's now see what is the answer for $a|1\rangle$ (what is your expectation?). For this, I'll calculate $\hat{n}a|1\rangle = 0$ because, again, $a^2 = 0$ (Pauli is so useful!). But we know that $\hat{n}|0\rangle = 0$, so we must have $a|1\rangle = C|0\rangle$ where C is some normalization constant, which is easily shown to be $C = 1$. So $a|1\rangle = |0\rangle$, *i.e.* removing a particle from the state with one particle, we get to the empty state (no particle). Now for the creation operator: $a^\dagger|0\rangle = a^\dagger a|1\rangle = \hat{n}|1\rangle = |1\rangle$, so adding a particle to the empty state gives us the state with one particle (no more comments needed, hopefully). And finally, $a^\dagger|1\rangle = (a^\dagger)^2|0\rangle = 0$, so indeed we cannot add a second particle if there is already one in that state - Pauli.

To summarize, after a bit of algebra we found that:

$$\begin{aligned} a^\dagger|0\rangle &= |1\rangle & a|0\rangle &= 0 \\ a^\dagger|1\rangle &= 0 & a|1\rangle &= |0\rangle \end{aligned}$$

so that we can write:

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

for the state with n particles in state α : we get to the state with n particles by adding n particles to the empty state – although for fermions $n = 0, 1$ only (we'll see soon why it is useful to put that $\sqrt{n!}$ there, even though it's always $= 1$ for fermions). This allows us now to generate any state $|n_1, n_2, \dots\rangle$ by acting with the appropriate creation operators the appropriate number of times – but before that, let's quickly catch up on bosons, as well.

Let's do the same for bosons, again considering a single one-particle state to simplify things. Note that its corresponding operators satisfy $[b, b^\dagger] = 1$ just like the raising and lowering operators for a single harmonic oscillator, so we should get the same eigenstates. Let me quickly remind you how this works: let $\hat{n}|n\rangle = n|n\rangle$ be an eigenstate of the number operator, then $n = \langle n|b^\dagger b|n\rangle \geq 0$. Using the commutation relations, one can show that $\hat{n}(b|n\rangle) = (n-1)(b|n\rangle)$ and $\hat{n}(b^\dagger|n\rangle) = (n+1)(b^\dagger|n\rangle)$, i.e. b and b^\dagger decrease or increase the eigenvalue by 1. In other words, if n is an eigenvalue, so is $n \pm 1, n \pm 2, \dots$. However, since we must have $n \geq 0$ for all possible eigenvalues, it follows that they must be non-negative integers, i.e. $n = 0, 1, 2, \dots$ and that we must have $b|0\rangle = 0$. The proper normalization is then found to be:

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

and again, one can express the state with n bosons in state α as:

$$|n\rangle = \frac{(b^\dagger)^n}{\sqrt{n!}}|0\rangle$$

for any $n = 0, 1, 2, \dots$. Just like for fermions, we find that b^\dagger increases the number of particles in that state (hence, creation operator), and b decreases them by one (hence, annihilation operator). Formally, this last equation looks just like what we obtained for fermions, so we can deal with both cases simultaneously.

Definition: we define a general Fock basis state as:

$$|n_1, n_2, \dots\rangle = \frac{(c_1^\dagger)^{n_1}}{\sqrt{n_1!}} \frac{(c_2^\dagger)^{n_2}}{\sqrt{n_2!}} \dots |0\rangle \quad (7)$$

where $|0\rangle = |0, 0, \dots\rangle$ is the vacuum. For bosons the n_i can have any value, for fermions $n_i = 0, 1$ only. **Note:** for fermions the order in which we write the operators is very important, because the operators anticommute: $a_1^\dagger a_2^\dagger = -a_2^\dagger a_1^\dagger$. This is why we have to define an order and stick to it!

The actions of the creation and annihilation operators on a Fock state can now be found by using the appropriate algebra. Let me show one example in detail:

$$c_\alpha| \dots, n_\alpha, \dots\rangle = c_\alpha \frac{(c_1^\dagger)^{n_1}}{\sqrt{n_1!}} \dots \frac{(c_\alpha^\dagger)^{n_\alpha}}{\sqrt{n_\alpha!}} \dots |0\rangle$$

c_α (anti)commutes with all operators before moving just before $(c_\alpha^\dagger)^{n_\alpha}$. There are $n_1 + \dots + n_{\alpha-1}$ such operators. We define the integers:

$$S_\alpha = \sum_{i=1}^{\alpha-1} n_i$$

and so:

$$c_\alpha \frac{(c_1^\dagger)^{n_1}}{\sqrt{n_1!}} \dots \frac{(c_\alpha^\dagger)^{n_\alpha}}{\sqrt{n_\alpha!}} \dots |0\rangle = \xi^{S_\alpha} \frac{(c_1^\dagger)^{n_1}}{\sqrt{n_1!}} \dots c_\alpha \frac{(c_\alpha^\dagger)^{n_\alpha}}{\sqrt{n_\alpha!}} \dots |0\rangle$$

Since (this is true for both fermions and bosons, see above)

$$c_\alpha |n_\alpha\rangle = \sqrt{n_\alpha} |n_{\alpha-1}\rangle$$

we find immediately that:

$$c_\alpha |n_1, \dots, n_\alpha, \dots\rangle = \xi^{S_\alpha} \sqrt{n_\alpha} |n_1, \dots, n_\alpha - 1, \dots\rangle$$

This looks similar to what we had above, except for the factor $\xi^{S_\alpha} = \pm 1$ which keeps track of the ordering and how many other states in “front” of state α are occupied. This sign is obviously very important for fermions (for bosons it is always 1). Similarly, one finds that:

$$c_\alpha^\dagger |n_1, \dots, n_\alpha, \dots\rangle = \xi^{S_\alpha} \sqrt{n_\alpha + 1} |n_1, \dots, n_\alpha + 1, \dots\rangle$$

So indeed, as advertised, creation operators add one more particle in the corresponding state, i.e. increase that occupation number by 1 (for fermions this is only allowed if $n_\alpha = 0$, of course, and the prefactor takes care of that); while annihilation operators remove one particle from that state. Because this is true for any state in the basis, it will also be true for any general states because they can be decomposed as a linear combination of basis states.

With these two equations, we can now compute the action of any operator because, as we show next, any operator can be written in terms of creation and annihilation operators.

Just one more example, to see that we get sensible results. Applying twice these rules, we find

$$c_\alpha^\dagger c_\alpha |n_1, \dots, n_\alpha, \dots\rangle = \xi^{S_\alpha} \sqrt{n_\alpha} c_\alpha^\dagger |n_1, \dots, n_\alpha - 1, \dots\rangle = \xi^{S_\alpha} \sqrt{n_\alpha} \xi^{S_\alpha} \sqrt{\xi(n_\alpha - 1) + 1} |n_1, \dots, n_\alpha, \dots\rangle$$

Since $\xi^2 = 1 \rightarrow \xi^{2S_\alpha} = 1$. If $\xi = 1 \rightarrow \sqrt{n_\alpha[\xi(n_\alpha - 1) + 1]} = n_\alpha$. If $\xi = -1$, then $\sqrt{n_\alpha[\xi(n_\alpha - 1) + 1]} = \sqrt{n_\alpha(2 - n_\alpha)} = n_\alpha$, because in this case we’re dealing with fermions, and we can only have $n_\alpha = 0, 1$. So we find, as expected, that:

$$\hat{n}_\alpha |n_1, \dots, n_\alpha, \dots\rangle = n_\alpha |n_1, \dots, n_\alpha, \dots\rangle$$

In other words, this operator indeed counts how many particles are in that state.

The only other question we need to answer now, is how to write general operators in terms of these creation and annihilation operators? Here’s the answer:

Theorem: If $\hat{A} = \sum_{i=1}^N \mathcal{A}_i$ is a single-particle operator, with \mathcal{A}_i acting only on particle “i”, then:

$$\hat{A} = \sum_{\alpha, \alpha'} \langle \alpha | \mathcal{A} | \alpha' \rangle c_\alpha^\dagger c_{\alpha'} \quad (8)$$

where

$$\langle \alpha | \mathcal{A} | \alpha' \rangle = \sum_{\sigma, \sigma'} \int d\vec{r} \phi_\alpha^*(\vec{r}, \sigma) \langle \vec{r}, \sigma | \mathcal{A} | \vec{r}, \sigma' \rangle \phi_{\alpha'}(\vec{r}, \sigma') \quad (9)$$

I will not reproduce the proof here because it’s long. We will discuss it in class and you can also find it in any standard textbook, for example Orland and Negele, or Fetter and Walecka. However, let me list the steps here. First, since a single-particle operator acts on just one particle (can be any of them), all it can do is either leave it in the state it was in, $\alpha \rightarrow \alpha$, or change its state $\alpha' \rightarrow \alpha$. All terms in Eq. (8) describe such processes. The question is how to find the coefficients for each process, namely $\langle \alpha | \mathcal{A} | \alpha' \rangle$. This is done so that the matrix elements of \hat{A} in a complete basis are correct – and it is a straightforward, but boring and time-consuming task to verify that indeed Eq.

(9) is the correct answer. Note: in writing Eq. (9), I assumed that the single-particle operator \mathcal{A} is diagonal in position, which is basically always the case. Before showing some examples, let me state the result for two-particle operators so that we're done with formalities:

Theorem: If $\hat{B} = \frac{1}{2} \sum_{i \neq j} \mathcal{B}_{i,j}$ is a two-particle operator, with $\mathcal{B}_{i,j}$ acting only on particles “i” and “j” ($1 \leq i, j \leq N$), then:

$$\hat{B} = \frac{1}{2} \sum_{\substack{\alpha, \alpha' \\ \beta, \beta'}} \langle \alpha, \beta | \mathcal{B} | \alpha', \beta' \rangle c_{\alpha}^{\dagger} c_{\beta}^{\dagger} c_{\beta'} c_{\alpha'} \quad (10)$$

where

$$\langle \alpha, \beta | \mathcal{B} | \alpha', \beta' \rangle = \sum_{\substack{\sigma_1, \sigma'_1 \\ \sigma_2, \sigma'_2}} \int d\vec{r}_1 \int d\vec{r}_2 \phi_{\alpha}^*(\vec{r}_1, \sigma_1) \phi_{\beta}^*(\vec{r}_2, \sigma_2) \langle \vec{r}_1, \sigma_1; \vec{r}_2, \sigma_2 | \mathcal{B} | \vec{r}_1, \sigma'_1; \vec{r}_2, \sigma'_2 \rangle \phi_{\alpha'}(\vec{r}_1, \sigma'_1) \phi_{\beta'}(\vec{r}_2, \sigma'_2) \quad (11)$$

NOTE the order of listing the annihilation operators in Eq. (10) !!! Do not list them in the “expected” order $c_{\alpha'} c_{\beta'}$ because for fermions that implies changing the sign of the interaction from attractive to repulsive (or vice-versa) and that's bound to make all the subsequent results very very wrong!

All I'm going to say about two-particle operators is that since they act on two particles, they can change the states of up to two particles, and that's precisely what all terms in Eq. (10) describe. And probably I'll give you the proof of this theorem as a homework, because it might be good to force you to calculate such an expectation value with wavefunctions, once in your lives. After that “enjoyable” experience, you'll adopt this 2nd quantization notation much more happily.

Let's see some examples, before you start thinking that this is all too complicated. We start first with some single-particle operators:

(a) The kinetic energy. In first quantization, the kinetic energy of a system with N particles is $\hat{T} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m}$. Since

$$\langle \vec{r}, \sigma | \frac{\hat{p}^2}{2m} | \vec{r}, \sigma' \rangle = \delta_{\sigma, \sigma'} \left(\frac{-\hbar^2}{2m} \nabla^2 \right)$$

(the well-known \vec{r} -representation of the kinetic energy of a single particle), in the 2ndQ we find:

$$\hat{T} = \sum_{\alpha, \alpha'} \langle \alpha | \frac{p^2}{2m} | \alpha' \rangle c_{\alpha}^{\dagger} c_{\alpha'}$$

where

$$\langle \alpha | \frac{p^2}{2m} | \alpha' \rangle = \sum_{\sigma} \int d\vec{r} \phi_{\alpha}^*(\vec{r}, \sigma) \left(\frac{-\hbar^2}{2m} \nabla^2 \right) \phi_{\alpha'}(\vec{r}, \sigma)$$

Let's go a bit farther, and assume that the basis $\alpha = (\vec{k}, \sigma)$ are the eigenstates of the momentum operator, so that $\phi_{\vec{k}, \sigma}(\vec{r}, \sigma) = \delta_{\sigma, \sigma'} \exp(i\vec{k} \cdot \vec{r}) / \sqrt{V}$ are just plane-waves. Then

$$\langle \vec{k}, \sigma | \frac{p^2}{2m} | \vec{k}', \sigma' \rangle = \delta_{\sigma, \sigma'} \delta_{\vec{k}, \vec{k}'} \frac{\hbar^2 \vec{k}^2}{2m}$$

(which you can either write down directly, or check by doing the integral I wrote above for these plane-waves basis states), and so, in this basis, the 2ndQ form of the kinetic energy is:

$$\hat{T} = \sum_{\vec{k}, \sigma} \frac{\hbar^2 \vec{k}^2}{2m} c_{\vec{k}, \sigma}^\dagger c_{\vec{k}, \sigma} \quad (12)$$

This, you should agree, makes a lot of sense. What this tells us is that this operator counts how many particles have a given momentum \vec{k} and spin σ (i.e., $\hat{n}_{\vec{k}, \sigma} = c_{\vec{k}, \sigma}^\dagger c_{\vec{k}, \sigma}$), multiplies that number by the kinetic energy associated with this state $\frac{\hbar^2 \vec{k}^2}{2m}$, and sums over all possible states. You should also check that the operator for the total number of particles $\hat{N} = \sum_{i=1}^N 1$ (in first quantization) $\rightarrow \sum_{\alpha} c_{\alpha}^\dagger c_{\alpha}$ (in second quantization), no matter what basis we use. Hopefully this also makes perfect sense to you!

Very important fact: the operator in the second-quantized form is independent of the number of particles N of the state it acts upon! In the first quantization, if we have a wavefunction with N particles, then the kinetic energy must be the sum of the N single-particle kinetic energies. If we have $N+2$ particles, then the kinetic energy is a different operator, with $N+2$ terms. By contrast, in the second-quantization the kinetic energy always is like in Eq. (12) (if we choose this basis). This means we can act with it on wavefunctions which are superpositions of states with different numbers of particles, no problem! This is a significant improvement, because the expression of the operators no longer changes if we change the number of particles – it is always the same!

I will show one more example, and give as assignment a few more usually encountered single-particle operators, so that you see that the final result is always what common-sense would predict.

(b) Total spin operator, $\vec{S} = \sum_{i=1}^N \vec{s}_i$. For simplicity, let us assume we're still in the basis $\alpha = (\vec{k}, \sigma)$. We need the matrix elements, let's compute them for each spin projection separately. I will assume that we work with spins $1/2$, which can be written in terms of the Pauli matrices $\vec{s} = \frac{\hbar}{2} \vec{\sigma}$. Then:

$$\langle \vec{k}, \sigma | s_z | \vec{k}', \sigma' \rangle = \frac{\hbar}{2} \delta_{\vec{k}, \vec{k}'} \delta_{\sigma, \sigma'} \sigma$$

so that

$$\hat{S}_z = \frac{\hbar}{2} \sum_{\vec{k}, \sigma} \sigma c_{\vec{k}, \sigma}^\dagger c_{\vec{k}, \sigma} = \frac{\hbar}{2} \sum_{\vec{k}} \left(c_{\vec{k}\uparrow}^\dagger c_{\vec{k}\uparrow} - c_{\vec{k}\downarrow}^\dagger c_{\vec{k}\downarrow} \right)$$

i.e, to get the z -axis spin one must subtract the number of particles with spin down from those with spin-up, and multiply by $\hbar/2$ – very reasonable, no? Similarly, we find:

$$\hat{S}_+ = \hat{S}_x + i\hat{S}_y = \hbar \sum_{\vec{k}} c_{\vec{k}\uparrow}^\dagger c_{\vec{k}\downarrow}$$

$$\hat{S}_- = \hat{S}_x - i\hat{S}_y = \hbar \sum_{\vec{k}} c_{\vec{k}\downarrow}^\dagger c_{\vec{k}\uparrow}$$

are indeed raising and lowering the total spin, by flipping spins-down into spins-up, or viceversa (while leaving the translational part, eg momentum carried by particles, unchanged). Putting all together, we can write:

$$\hat{\vec{S}} = \sum_{\vec{k}, \sigma, \sigma'} c_{\vec{k}, \sigma}^\dagger \frac{\hbar \vec{\sigma}_{\sigma\sigma'}}{2} c_{\vec{k}, \sigma'}$$

where $\vec{\sigma}_{\sigma\sigma'}$ are the matrix elements of the Pauli matrices. Again, this 2ndQ expressions are valid no matter how many particles are in the system!

(c) Two-particle operator $\hat{V} = \frac{1}{2} \sum_{i \neq j} u(\vec{r}_i - \vec{r}_j)$ for two-particle interactions (for instance, $u(\vec{r}) = e^2/r$ can be the Coulomb repulsions between electrons). For simplicity, I will assume that the interactions depend only on the distance between particles, but not on their spins (true for Coulomb and all other examples we'll consider in this course. But more general options exist, and you can deal with them using Theorem 2).

We'll continue to work in the basis $\alpha = (\vec{k}, \sigma)$, so that the basis wavefunctions are simple planewaves $\langle \vec{r} | \vec{k} \rangle = e^{i\vec{k}\vec{r}}/\sqrt{V}$, where V is the volume of the system. (I will assume V finite to begin with, since in this case the allowed \vec{k} values are discrete, and it makes sense to write $\sum_{\vec{k}}$, not $\int d\vec{k}$. We can let the volume $V \rightarrow \infty$ at the end, if need be. We'll go more carefully through the details of this procedure a bit later on, when we'll need to use it). Because:

$$\langle \vec{r}_1, \sigma_1; \vec{r}_2, \sigma_2 | u(\hat{\vec{r}} - \hat{\vec{r}}') | \vec{r}_1, \sigma'_1; \vec{r}_2, \sigma'_2 \rangle = u(\vec{r}_1 - \vec{r}_2) \delta_{\sigma_1, \sigma'_1} \delta_{\sigma_2, \sigma'_2}$$

Eq. (11) becomes:

$$\langle \vec{k}_1, \sigma_1; \vec{k}_2, \sigma_2 | u | \vec{k}_3, \sigma_3; \vec{k}_4, \sigma_4 \rangle = \delta_{\sigma_1, \sigma_3} \delta_{\sigma_2, \sigma_4} \int d\vec{r}_1 \int d\vec{r}_2 \frac{e^{-i\vec{k}_1\vec{r}_1 - i\vec{k}_2\vec{r}_2}}{V} u(\vec{r}_1 - \vec{r}_2) \frac{e^{i\vec{k}_3\vec{r}_1 + i\vec{k}_4\vec{r}_2}}{V}$$

We now define the Fourier transform of the interaction potential:

$$u(\vec{r}) = \frac{1}{V} \sum_{\vec{q}} e^{i\vec{q}\vec{r}} u_{\vec{q}}; \quad u_{\vec{q}} = \int d\vec{r} e^{-i\vec{q}\vec{r}} u(\vec{r})$$

Using the identity $\int d\vec{r} e^{i(\vec{k} - \vec{k}')\vec{r}} = V \delta_{\vec{k}, \vec{k}'}$ we can now do the two integrals over \vec{r}_1, \vec{r}_2 easily. The end result, which we will use extensively, is:

$$\hat{V} = \frac{1}{2V} \sum_{\substack{\vec{k}, \vec{k}' \\ \sigma, \sigma'}} \sum_{\vec{q}} u_{\vec{q}} c_{\vec{k} + \vec{q}, \sigma}^\dagger c_{\vec{k}' - \vec{q}, \sigma'}^\dagger c_{\vec{k}', \sigma'} c_{\vec{k}, \sigma} \quad (13)$$

In other words, two particles that initially were in the states $(\vec{k}, \sigma), (\vec{k}', \sigma')$ have interacted with one another, exchanged some momentum \vec{q} (but with total momentum conserved, as it should be since this interaction is invariant to global translations), and end up in the states $(\vec{k} + \vec{q}, \sigma), (\vec{k}' - \vec{q}, \sigma')$. The vertex $u_{\vec{q}}$ is directly determined by $u(\vec{r})$ and basically shows which \vec{q} is more likely to be exchanged during the interaction. Because this two-particle interaction does not act on spins, these remain unchanged between the initial and final states. Of course, if we choose a more complicated two-particle interaction that also involves the spins somehow, then its second quantization expression will reflect that. In this course we will only consider Coulomb interactions which do not depend on spins, so the equation listed above is the only one we will need.

We can similarly write the second-quantization expression for any operator, in terms of creation and annihilation operators. Because we know how c, c^\dagger act on the Fock basis states $|n_1, \dots\rangle$, we can now easily and efficiently perform calculations such as finding matrix elements of various operators. In particular, consider how much easier it is now to calculate $\langle \Psi | \hat{O} | \Psi \rangle$ for some $|\Psi\rangle = \sum_{n_1, n_2, \dots} c_{n_1, n_2, \dots} |n_1, n_2, \dots\rangle$, i.e. if we work in the abstract space using the 2nd quantization (contrast this with the discussion at the bottom of page 4, for wavefunctions). Unlike there, here

we don't have to perform the N integrals over positions, the $2N$ sums over spins, and moreover each $|n_1, n_2, \dots\rangle$ is a single object (an abstract vector), not a (antisymmetrized) sum of $N!$ products of one-particle wavefunctions.

So doing calculations is much simpler, once we adopt the second quantization notation, and in plus we are no longer forced to work with states that have a fixed number of particles, we can easily deal with grand-canonical ensembles. You might say, though, that having to list an infinite number of occupation numbers n_1, n_2, \dots for each basis state is less than ideal. And you are right, but in practice we don't do that. Let me denote the state without particles $|n_1 = 0, n_2 = 0, \dots\rangle \equiv |0\rangle$ (we could call this the vacuum, but don't forget that it's only the vacuum for the particles we're describing as possibly occupying these one-particle states – example, it could be the vacuum for the valence electrons, but that still leaves lots of core electrons and nuclei in the system). This $|0\rangle$ is the only state with $N = 0$. Consider now states with $N = 1$, i.e. with one particle. If the particle occupies state α , this is the state $|\dots, 0, 1, 0, \dots\rangle$ where the 1 corresponds to $n_\alpha = 1$, and all other $n_\beta = 0$. As discussed, this state is generated from the vacuum as $|\dots, 0, 1, 0, \dots\rangle = c_\alpha^\dagger |0\rangle$. Similarly, a state with one particle in state α and one in $\beta \neq \alpha$ is $|\dots, 0, 1, 0, \dots, 0, 1, 0, \dots\rangle = c_\alpha^\dagger c_\beta^\dagger |0\rangle$. In practice, we'll use the rhs notation for states, which is more compact than the lhs one. In other words, we don't bother listing all the empty states, we only specify which ones are occupied. For fermions, we need to be very careful with the order of the operators, because $c_\alpha^\dagger c_\beta^\dagger |0\rangle = -c_\beta^\dagger c_\alpha^\dagger |0\rangle$ etc. Those signs can wreak havoc if we're not carefully keeping track of them, as we'll see in some examples in class and/or homeworks.

Before moving on, let me comment on a special one-particle basis, namely $\alpha = (\vec{r}, \sigma)$ – the basis associated with the position and spin operators. Can we use this basis for the 2nd quantization? The answer is of course yes, because it is a complete one-particle basis set, and that is all we asked for. It is customary to use special notation for the creation and annihilation operators associated with adding/removing a particle with spin σ from position \vec{r} , namely $\hat{\Psi}_\sigma^\dagger(\vec{r}), \hat{\Psi}_\sigma(\vec{r})$ (instead of $c_{\vec{r},\sigma}^\dagger, c_{\vec{r},\sigma}$, although they are precisely the same objects), and to call them **field operators**. We will not need these operators in this course, so I will not insist on this topic. However, you will need them in Phys503 and in general, so it is useful to at least have some idea how they work (which, really, is just like any other set of creation and annihilation operators). I added some supplementary material at the end of this section if you want to read more about them. I also explain there why this notation is called **the second quantization**.

4 Examples of well-known Hamiltonians

4.1 The Jellium Model

This model describes a system of N interacting valence electrons, which are placed within a uniform background of positive charge, so that the whole system is charge neutral. In other words, it is as if we “smear out” the ions into a uniform distribution of positive charge occupying the whole volume. This greatly simplifies things because the resulting system is invariant to all translations, instead of being invariant only to lattice translations. Note that this also means that we now only need to consider the motion of the valence electrons within this inert background of positive charge (i.e., we no longer need to worry about lattice vibrations and their influence on the valence electrons, because there is no lattice as such). So within the jellium model, only the behavior of the valence electrons needs to be understood.

Before continuing the discussion, let's see why in the world would we want to pretend that the positive ions are smushed into a continuous uniform background. The answer is that sometimes

this is a very good approximation (and one that, as we will see, will make calculations much easier because it increases the symmetry of the system). The valence electrons have a characteristic lengthscale λ_F (the Fermi wavelength – you should know about this from an undergrad stat mech course, but if not, we'll review soon what this is). The lattice, on the other hand, is characterized by a lattice constant a . In cases where $\lambda_F \gg a$, it becomes a good approximation to set $a = 0$, i.e. to pretend that there is no discrete lattice, just a continuous positive charge distribution. By the way, our brains make this approximation all the time. For example, when we walk, we know very well that the floor is made out of atoms separated by some characteristic distance (the a) but that distance is so small compared to the size of our foot (the analog of λ_F , ie our lengthscale) that we can pretend that the floor is continuous. If we were much tinier so that the size of our feet was comparable to this distance $\lambda_F \sim a$, then we would be very careful where we step, so that we don't fall into the space between two atoms. In that case the approximation that the floor is continuous would be terrible. The same is true for our electrons. If they happen to be spread over $\lambda_F \gg a$, then the jellium model is a very good descriptions. In other cases, we will find that $\lambda_F \sim a$, in which case we need to use models (such as the Hubbard model, described next) which don't make this jellium approximation.

Because there is no lattice, the jellium model has the full symmetry to all translations (not just a discrete set, like for a lattice), and the best one-particle basis are plane-waves with the quantum numbers $\alpha = (\vec{k}, \sigma)$ (note: from now on r, p, k, \dots are all understood to be vectors, but I will no longer put arrows on them, except in the final expressions).

In the first quantization, the Hamiltonian for this system is:

$$\hat{H} = \hat{H}_{el} + \hat{H}_{backg} + \hat{H}_{el-backg}$$

where the electronic Hamiltonian is:

$$\hat{H}_{el} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} u(r_i - r_j) \quad (14)$$

where we will assume a so-called screened Coulomb interaction between electrons

$$u(r) = \frac{e^2}{r} e^{-\mu r}$$

(at the end we can take $\mu \rightarrow 0$, if we want the usual, unscreened Coulomb interaction).

The background has a uniform positive density $n(r) = N/V$, and its energy is:

$$\hat{H}_{backg} = \frac{1}{2} \int dr \int dr' n(r) u(r - r') n(r') = \frac{N^2}{2V} u_{q=0}$$

since $\int dr = V$. From now on, I will denote the $q = 0$ Fourier component of the interaction as u_0 . Since $u_q = \int dr e^{-i\vec{q}\vec{r}} u(r) = 4\pi e^2 / (q^2 + \mu^2) \rightarrow u_0 = 4\pi e^2 / \mu^2$, and you see that we need a finite μ , at least for the time being.

The electron-background interactions are described by:

$$\hat{H}_{el-backg} = - \sum_{i=1}^N \int dr n(r) u(r - r_i) = - \frac{N^2}{V} u_0$$

because $\sum_i 1 = N$. So things simplify considerably because of the “jellium” assumption and we find:

$$\hat{H} = \sum_{i=1}^N \frac{\hat{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} u(r_i - r_j) - \frac{N^2}{2V} u_0$$

We now go to the second quantization, and use the (k, σ) basis, to find:

$$\hat{H} = \sum_{k, \sigma} \frac{\hbar^2 k^2}{2m} c_{k\sigma}^\dagger c_{k\sigma} + \frac{1}{2V} \sum_{\substack{k, k' \\ \sigma, \sigma'}} \sum_q u_q c_{k+q, \sigma}^\dagger c_{k'-q, \sigma'}^\dagger c_{k', \sigma'} c_{k, \sigma} - \frac{N^2}{2V} u_0$$

(these terms were discussed individually in the previous section). Let us consider the $q = 0$ contribution from the electron-electron interaction:

$$\frac{u_0}{2V} \sum_{\substack{k, k' \\ \sigma, \sigma'}} c_{k, \sigma}^\dagger c_{k', \sigma'}^\dagger c_{k', \sigma'} c_{k, \sigma} = \frac{u_0}{2V} \sum_{\substack{k, k' \\ \sigma, \sigma'}} c_{k, \sigma}^\dagger c_{k\sigma} c_{k', \sigma'}^\dagger c_{k', \sigma'} = \frac{u_0}{2V} N^2$$

Here, in the second equality we use the fact that $(k\sigma) \neq (k', \sigma')$ because if they are equal, $c_{k\sigma} c_{k\sigma} = 0$ (electrons are definitely fermions). As a result, we can anticommute $c_{k\sigma}$ past the two other operators to obtain the middle expression. In the last equality, we used the fact that $\sum_{k\sigma} c_{k, \sigma}^\dagger c_{k, \sigma} = N$ if the system contains N electrons (strictly speaking, the condition $(k\sigma) \neq (k', \sigma')$ means that we have $N(N-1)$ instead of N^2 , but in the thermodynamic limit $N \rightarrow \infty$, this makes no difference). So we see that this exactly cancels the term left from the backg. and el-backg Hamiltonians (so we are no longer troubled even if $\mu \rightarrow 0$), and we obtain **the jellium model Hamiltonian**:

$$\hat{H} = \sum_{\vec{k}, \sigma} \frac{\hbar^2 k^2}{2m} c_{\vec{k}\sigma}^\dagger c_{\vec{k}\sigma} + \frac{1}{2V} \sum_{\substack{\vec{k}, k' \\ \sigma, \sigma'}} \sum_{\vec{q} \neq 0} u_{\vec{q}} c_{\vec{k}+\vec{q}, \sigma}^\dagger c_{\vec{k}'-\vec{q}, \sigma'}^\dagger c_{\vec{k}', \sigma'} c_{\vec{k}, \sigma} \quad (15)$$

Note: this looks the same whatever the interaction happens to be (so long as it is spin independent, which again, is the only kind of interaction we'll consider in this course). For instance, we may want to consider short-range interactions $u(\vec{r}) = U\delta(\vec{r}) \rightarrow u_{\vec{q}} = U$ instead of the Coulomb interaction. The Hamiltonian has the same form, we just have to use the appropriate u_q .

4.2 The Hubbard model

This works well in the opposite case to the one discussed above, where the electronic wavefunctions extend over small distances compared to a . In this case, we start with atoms placed on a regular lattice, and use a basis of Wannier orbitals as our one-particle basis. (Wannier orbitals are eigenfunctions which are strongly localized about individual lattice sites. You can think of them almost as being the atomic orbitals of the isolated atom – except that some corrections are needed to insure that they are all orthogonal to one another and the basis is complete. The details can be rather complicated, but I hope it's reasonable to believe that such a basis exists). In this case, the most suitable one-particle states basis is identified by $\alpha = (i, n, \sigma)$, where i indicates the lattice site where the orbital is localized (spatial location \vec{R}_i), n is a shorthand notation for the set of quantum numbers identifying the type of orbital (for instance, $1s$ or $2p_z$) and σ is the spin. Then

$$\langle \vec{r}, \sigma | i, n, \sigma' \rangle = \delta_{\sigma, \sigma'} \phi_n(\vec{r} - \vec{R}_i)$$

is the atomic-like wavefunction associated with this state. We then define $a_{i,n,\sigma}^\dagger, a_{i,n,\sigma}$ as the associated creation and annihilation operators for electrons with the spin σ , into the orbital n of the i^{th} atom.

Let us start again from our CM “Theory of everything” + BOA, but let’s not make the ions vs. valence electrons approximation yet. In other words, we have the nuclei frozen at their equilibrium lattice positions, and all the electrons (both core and valence) occupying various orbitals about various nuclei. We’ll see at the end why/when we can ignore the core electrons and return to thinking about ions + valence electrons only.

With this approximation, the Hamiltonian in the second quantization becomes:

$$\mathcal{H} = \sum_{\substack{i,n,\sigma \\ j,m,\sigma'}} \langle i,n,\sigma | \hat{T} + \hat{V}_{ext} | j,m,\sigma' \rangle a_{i,n,\sigma}^\dagger a_{j,m,\sigma'} \\ + \frac{1}{2} \sum_{\substack{i_1,n_1,\sigma_1,i'_1,n'_1,\sigma'_1 \\ i_2,n_2,\sigma_2,i'_2,n'_2,\sigma'_2}} \langle i_1,n_1,\sigma_1; i_2,n_2,\sigma_2 | \hat{V}_{e-e} | i'_1,n'_1,\sigma'_1; i'_2,n'_2,\sigma'_2 \rangle a_{i_1,n_1,\sigma_1}^\dagger a_{i_2,n_2,\sigma_2}^\dagger a_{i'_2,n'_2,\sigma'_2} a_{i'_1,n'_1,\sigma'_1}$$

The first line has the contribution from one-particle operators, i.e. the kinetic energy $\hat{T} = \hat{p}^2/2m$ and the interaction of the electrons with the potential created by the frozen ions $\hat{V}_{ext} = -Z \sum_i v(\vec{r} - \vec{R}_i)$; and the second one has the two-particle operators, i.e. the electron-electron repulsion $\hat{V}_{e-e} = v(\vec{r} - \vec{r}')$, where $v(r)$ is the (maybe screened) Coulomb interaction. Of course, there is also a contribution from the nucleus-nucleus repulsion, but if the nuclei are frozen that is just some overall constant and I am not writing it here (that constant is very important for the cohesion energy of the material, but it’s not affecting what the electrons do).

Let us consider $(\hat{T} + \hat{V}_{ext})|j,m,\sigma'\rangle$ first. We can divide the potential into the term j corresponding to this particular nucleus plus the interaction with all other nuclei: $\hat{V}_{ext} = -Zv(\vec{r} - \vec{R}_j) - Z \sum_{l \neq j} v(\vec{r} - \vec{R}_l)$. If $|j,m,\sigma'\rangle$ is an atomic orbital, then $(\hat{T} - Zv(\vec{r} - \vec{R}_j))|j,m,\sigma'\rangle = E_m|j,m,\sigma'\rangle$ i.e. it is an eigenstate of the one-particle problem with only that one nucleus in the system. Here, E_m are the atomic energies of the $1s, 2s, \dots$ corresponding levels. As a result, we can write the first line as:

$$\mathcal{H}_1 = \sum_{i,n,\sigma} E_n a_{i,n,\sigma}^\dagger a_{i,n,\sigma} - \sum_{\substack{i,n,\sigma \\ j,m}} t_{i,n;j,m} a_{i,n,\sigma}^\dagger a_{j,m,\sigma}$$

where

$$t_{i,n;j,m} = \langle i,n,\sigma | Z \sum_{l \neq j} v(\vec{r} - \vec{R}_l) | j,m,\sigma \rangle$$

is due to the attraction that the electron in the state j,m,σ feels from all other nuclei $l \neq j$ in the system. Note that this term is diagonal in spin. So looking at \mathcal{H}_1 , the first term counts the energy of the electrons if they were in isolated atoms – how many are on each level, times their atomic level energy. The second reflects the influence of the other nuclei. The second term can be split into two parts: there are terms where $i = j, n = m$, namely $-\sum_{i,n,\sigma} t_{i,n;i,n} a_{i,n,\sigma}^\dagger a_{i,n,\sigma}$. By symmetry, $t_{i,n;i,n} \equiv t_n$ cannot depend on which nucleus we’re talking about. So these can be combined with the atomic terms and change $E_n \rightarrow E_n - t_n = \tilde{E}_n$. This shows that one effect of the other nuclei is to lower the value of the effective atomic energies, simply because an electron feels attraction not just from its own nucleus but there is also some negative potential created by all other nuclei in the system. Depending on crystal structure, this additional potential may lift

some of the degeneracies that are present in the isolated atom, for example between the p_x, p_y, p_z orbitals, or between the $d_{xy}, d_{xz}, d_{yz}, d_{x^2-y^2}, d_{3z^2-r^2}$. This is known as **crystal field** effects. In class we'll briefly discuss a simple example of how the degeneracy between a d_{xy} and a $d_{x^2-y^2}$ orbital is lifted in a square lattice.

The remaining terms change the state of the electron because they move it from one orbital into a different one. There are terms with $i = j$ but $n \neq m$, which keep the electron in the same atom but change its orbital. These are always ignored, so far as I know. One possible explanation is as follows. Consider the matrix element $\langle i, n, \sigma | Z \sum_{l \neq i} v(\vec{r} - \vec{R}_l) | i, m, \sigma \rangle$ associated with such terms. The lattice is symmetric (eg, cubic) so the total potential created by all other nuclei is almost spherical, because the other nuclei are placed very orderly around the site i . If the potential was truly spherical, then angular momentum conservation would guarantee that these terms vanish for any orbitals n, m with different angular momentum quantum numbers. So how good or bad is this approximation is a matter of how close or far is this potential from being spherically symmetric. For inner orbitals that are located very close to nucleus i , the approximation is very good. For more extended orbitals the approximation is more and more dubious. Depending on the symmetry of the crystal and if one uses the so-called real wavefunctions, i.e. p_x, p_y, p_z instead of $\phi_{n,l=1,m}(\vec{r})$ with $m = -1, 0, 1$ and $d_{xy}, d_{xz}, d_{yz}, d_{x^2-y^2}, d_{3z^2-r^2}$ instead of $\phi_{n,l=2,m}(\vec{r})$ with $m = -2, \dots, 2$, additional symmetries may also guarantee that these off-diagonal matrix elements vanish. In any event, such terms are ignored and we will do so as well in the following discussion.

Finally, we also have terms with $i \neq j$, which show that due to attraction from other nuclei, an electron can “hop” from one orbital of one atom into another orbital of another atom. The energies $t_{i,n,j,m}$ associated with this are called **hopping matrices** or **hopping integrals**. Again, it is customary to make some approximations and not keep all such terms. A very reasonable one comes from keeping only terms where i and j are nearest-neighbor sites, because the atomic orbital wavefunctions decay exponentially so these terms will be the largest of all. Not surprisingly, this is called the **nearest-neighbor hopping model**. (Of course, when necessary we can include longer range hopping as well). Because the nuclei are on a lattice, all these matrix elements have equal magnitude although the sign may change, depending on the sign of the orbitals involved. To be more specific, let i and j be two adjacent sites, and let me approximate:

$$t_{i,n;j,m} \approx \langle i, n, \sigma | Z v(\vec{r} - \vec{R}_i) | j, m, \sigma \rangle = Z \int d\vec{r} \phi_n^*(\vec{r} - \vec{R}_i) v(\vec{r} - \vec{R}_i) \phi_m(\vec{r} - \vec{R}_j)$$

where I kept from $\sum_{l \neq j}$ only the term $l = i$ which will contribute most to the matrix element (all other nuclei are even farther away so they create a much smaller potential). Given that any pair of nearest-neighbor nuclei are at the same distance $\vec{R}_i - \vec{R}_j$ from each other, the result cannot depend on this distance and must have the same magnitude for all such hoppings.

The sign, however, can change depending on details. By definition $v(r)$ is a positive quantity, so the sign of this integral will depend on the signs of the orbitals. If they are both $1s$ -type then they are positive everywhere and the corresponding t is positive and the same for all pairs of neighbor atoms. But if one of them is s -type and the other is a p -type orbital, then the question is whether the p lobe pointing towards the s -orbital is positive or negative – that will decide the sign of the matrix element. If you think about it, if we consider hopping from a p -orbital into orbitals of atoms to its right and its left, those will have different signs (but the same magnitude) because one integral is controlled mostly by the negative lobe, and one by the positive one. You can also see that symmetry will set some of these hoppings to zero, for instance if we have two atoms along the x axis, there will be zero hopping between s orbitals of one atom and p_y and p_z orbitals of the other one. And so on and so forth. We'll discuss several examples in class, because this is an

important point.

To summarize what we have so far: the single-particle operators contribution to the total Hamiltonian is:

$$\mathcal{H}_1 = \sum_{i,n,\sigma} \tilde{E}_n a_{i,n,\sigma}^\dagger a_{i,n,\sigma} - \sum_{\langle i,j \rangle, n, m, \sigma} t_{i,n;j,m} a_{i,n,\sigma}^\dagger a_{j,m,\sigma}$$

where the notation $\langle i, j \rangle$ is used to show that the sum is only over pairs of nearest neighbor sites. If we stop here (i.e., ignore electron-electron interactions), we have a hopping Hamiltonian, as I said. As is the case for any model Hamiltonian, we don't know the proper values for its parameters $\tilde{E}_n, t_{i,n;j,m}$ although as discussed above, we know the signs of the hoppings. Usually these quantities are used as free parameters and their values are adjusted so as to obtain agreement with experimentally measured quantities.

Now consider the two-particle contribution, due to electron-electron repulsion. The general expressions shows that this repulsion, when acting between electrons occupying orbitals i_2, n_2 and i'_n, n'_2 , can scatter them into the orbitals i_1, n_1 and i'_1, n'_1 . Because the repulsion does not depend on spin, the spins of the electrons remain unchanged: $\sigma_1 = \sigma'_1, \sigma_2 = \sigma'_2$. Again, some of these terms will be considerably bigger than others, and in fact you can easily see that the largest terms are when $i_1 = i'_1 = i_2 = i'_2$, because then all 4 atomic orbitals are located at the same site so their overlap (which controls the size of the matrix element) is as big as possible. Physically, this makes good sense: the strongest interactions must be between electrons belonging to the same nucleus, because they are closest together. If we make the approximation that we only keep these terms, we say that we only keep **on-site repulsions**. Again, if needed, we can also add repulsion between nearest-neighbors and even longer-range ones.

Even if we keep only on-site repulsion, we could still scatter two electrons from any two orbitals of any site into any two other orbitals of the same site, so there is still a huge number of possibilities. So let me now divide the electrons into core electrons and valence electrons.

For the core electrons, which are very close to the nucleus, $\tilde{E}_n \approx E_n$ and $t_{in,jm} \approx 0$, because corrections from other nuclei are tiny compared to the very strong potential of its own nucleus. We can also ignore the contributions from the electron-electron repulsion term, for the following reason: because all core states are full, repulsion can only scatter two core electrons into two empty levels (Pauli principle). But those are located at energies that are much higher, and this difference in energy is considerably bigger than the typical repulsion matrix element. Roughly put, the repulsion is not strong enough so the probability of such processes is very small and can safely be ignored. So you can see that we arrived at the conclusion that the core electrons basically are not influenced by the presence of the other atoms, they just sit on their atomic levels of energy E_n and are inert, they do not hop to other atoms or be scattered because of electron-electron repulsion (similar arguments offer a second explanation as to why we can ignore the single-particle hopping between different orbitals of the same atom). Which is why we can leave out the core electrons altogether, we already know precisely what they're doing and we do not need to make any further calculations concerning them.

For the valence electrons, however, the hopping integrals cannot be ignored (remember that these are the most spatially extended orbitals, so they have the largest hopping integrals). We also cannot ignore their Coulomb repulsions, because these orbitals are usually only partially filled, so there are empty orbitals with the same energy into which electrons can be scattered by repulsion.

For simplicity, let me assume that the valence electrons are in a partially-filled s -type orbital. Then, there is a single on-site repulsion matrix element associated with scattering, i.e.:

$$U = \langle i, s; i, s | V_{e-e} | i, s; i, s \rangle = \int d\vec{r} \int d\vec{r}' |\phi_s(\vec{r})|^2 |\phi_s(\vec{r}')|^2 v(\vec{r} - \vec{r}')$$

Of course, these valence electrons could also be scattered from their s orbital into higher, empty orbitals (the lower ones are all filled with core electrons). But if those are at much higher energies than the typical repulsion matrix element, such processes are again very unlikely to occur and are ignored. So in this case, all that is left of the two-particle repulsion term is:

$$\mathcal{H}_2 = \frac{U}{2} \sum_{i,\sigma,\sigma'} a_{i\sigma}^\dagger a_{i\sigma'}^\dagger a_{i\sigma'} a_{i\sigma} = U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow}$$

The second equality follows because we must have $\sigma \neq \sigma'$, otherwise $a_{i\sigma} a_{i\sigma} = 0$. We can then rearrange that term by commuting $a_{i\sigma}$ past the σ' operators, and rewriting it in terms of $\hat{n}_{i\sigma} = a_{i\sigma}^\dagger a_{i\sigma}$ (we assume spins 1/2). This is known as the on-site Hubbard repulsion, and if we add the hopping terms (only for these valence electrons), we get **the Hubbard Hamiltonian**

$$\hat{H} = -t \sum_{\langle i,j \rangle, \sigma} a_{i\sigma}^\dagger a_{j\sigma} + U \sum_i \hat{n}_{i\uparrow} \hat{n}_{i\downarrow} \quad (16)$$

Here I didn't bother to specify anymore that the valence orbitals are s , because they are now the only orbitals left in the problem. I also used the fact that for s orbitals, all nearest-neighbor hopping have the same sign $t > 0$.

Of course, this model can be extended to include hopping to second-nearest neighbors etc, and also longer-range interaction terms, between atoms which are on neighbor sites. Another type of generalization is to multiple orbitals. If, for example, we have a Wannier orbital of p -type, i.e. three-degenerate, then we would have 6 valence states associated with each atom (3 orbitals x 2 spin projections), and if these orbitals remain degenerate in the solid then we would need to keep all of them in the Hamiltonian. In this case, we have extra indices to indicate which of the three orbitals is involved, and hopping will strongly depend on the combination of orbitals involved, as discussed. The on-site interaction terms also become much more complicated because electrons can scatter from any 2 of these orbitals into any 2 others (inside the same atom), so more terms appear. They can be calculated based on the symmetries of the orbitals and they are tabulated in books. So things can quickly become very complicated, it all depends on the complexity of the material we are trying to describe. If we have multiple types of atoms, then that adds an extra layer of complexity, but if you think about it modelling will proceed in many ways in somewhat similar terms: we have to figure out which are the valence electrons, what sort of hopping they can have and what sort of on-site interactions they can have, consistent with the symmetries of the problem at hand. That gives us the simplest model Hamiltonian. If it works well (ie it its predictions agrees with experimental measurements) then great, the problem is solved. If not, we have to go back, figure out what are the next biggest terms we ignored, include them and keep trying until we get the “right” model. So you can see how this is less pleasant than dealing with ab-initio models, where there is no such guessing because we know the Hamiltonian with all its parameters; but you can also see that these model Hamiltonians are much simpler than the ab-initio ones, so we have some chance to get some answers.

To put things in perspective, let me mention that even the Hubbard Hamiltonian of Eq. (16) is not yet understood (except in some particular cases), despite around 50 years of intense study!

Before moving on, let me also say that the “derivation” I offered here, while more detailed than what is typically done in textbooks, is still sweeping many things under the carpet. For instance, you might wonder why the hopping integrals are controlled by the bare electron-nucleus potential (proportional to Z) and not the screened one (proportional to Z^*); in other words, why are the core electrons not screening this potential? In fact they do, if you track carefully the effect of some of

the terms I conveniently “forgot” to mention - for instance, what happens when a valence electron scatters on a core electron. The core electron can stay where it was and the valence electron can be scattered to another atom, so this looks like an effective hopping, and indeed you should convince yourself that such terms will screen the potential. But the bottom line is that we do not calculate the value of these hopping integrals anyway, they are free parameters, so what matters is to figure out what are the largest contributions possible (hoppings and repulsions) and what restrictions must be imposed due to symmetries. That is enough to set up a model Hamiltonian. We’ll discuss some more examples in class, if we have time.

4.3 Liquid of He^4 atoms

Consider a collection of He^4 atoms (these are bosons, unlike He^3). Because they are so light, they remain liquid at all temperatures at normal pressures, so in this case a Born-Oppenheimer approximation makes no sense. We can study the properties of such a quantum liquid with the formalism we have discussed. For simplicity, we assume that each He holds on to its electrons, i.e. we study the properties of the atoms, not ions + electrons. In first quantization, the corresponding Hamiltonian is:

$$\hat{H} = \sum_{i=1}^N \frac{\vec{P}_i^2}{2M} + \frac{1}{2} \sum_{i \neq j} V(\vec{R}_i - \vec{R}_j) \quad (17)$$

where $V(\vec{r})$ is some short-range interactions between these atoms.

A liquid is invariant to all translations so it makes sense again to use \vec{k} -basis plane-waves, and we don’t need a spin index because the spin is zero. Then, following all the rules we derived so far (but now for bosonic operators) we find that in the 2nd quantization:

$$\hat{H} = \sum_{\vec{k}} \frac{\hbar^2 \vec{k}^2}{2M} b_{\vec{k}}^\dagger b_{\vec{k}} + \frac{1}{2V} \sum_{\vec{k}, \vec{k}', \vec{q}} V_{\vec{q}} b_{\vec{k}+\vec{q}}^\dagger b_{\vec{k}'-\vec{q}}^\dagger b_{\vec{k}'} b_{\vec{k}} \quad (18)$$

where, again, $V_{\vec{q}}$ is the Fourier transform of $V(\vec{r})$. This looks formally very similar to the jellium Hamiltonian that describes a liquid of electrons, apart from the fact that those operators were fermionic, while these are bosonic. This is another nice feature of 2nd quantization, many equations look similar for fermions and bosons so one does not need to remember twice as many formulae.

Hopefully these examples gave you a good enough idea of how we can obtain various model Hamiltonians, and what are their general expressions in second quantization notation. As advertised before, we will next discuss how to solve purely electronic Hamiltonians, and then move on to the other steps discussed at the end of the previous section.

5 Supplementary material: field operators

As promised, let us discuss what happens if we choose the one-particle basis (\vec{r}, σ) . This, of course, can be done (or I would not have mentioned it). We define the *field operators*:

$$\hat{\Psi}(\vec{r}) = \sum_{\alpha} \phi_{\alpha}(\vec{r}) c_{\alpha}; \quad \hat{\Psi}^\dagger(\vec{r}) = \sum_{\alpha} \phi_{\alpha}^T(\vec{r}) c_{\alpha}^\dagger \quad (19)$$

where (see the first section)

$$\phi_\alpha(\vec{r}) = \begin{pmatrix} \phi_\alpha(\vec{r}, +S) \\ \dots \\ \phi_\alpha(\vec{r}, -S) \end{pmatrix} \text{ and } \phi_\alpha^T(\vec{r}) = (\phi_\alpha^*(\vec{r}, +S), \dots, \phi_\alpha^*(\vec{r}, -S))$$

are $2S + 1$ -spinors. Using the completeness of the basis α (this can be any one-particle basis we like, eg (k, σ)) and the commutation relations $[c_\alpha^\dagger, c_\beta]_\xi = \delta_{\alpha, \beta}$ etc., it is straightforward to show that the field operators satisfy the proper commutation relations like any other choice of creation and annihilation operators:

$$[\hat{\Psi}(\vec{r}), \hat{\Psi}^\dagger(\vec{r}')]_\xi = \delta(\vec{r} - \vec{r}'); \quad [\hat{\Psi}(\vec{r}), \hat{\Psi}(\vec{r}')]_\xi = [\hat{\Psi}^\dagger(\vec{r}), \hat{\Psi}^\dagger(\vec{r}')]_\xi = 0 \quad (20)$$

Before going on, let's see what is the meaning of these operators. Consider the action of $\hat{\Psi}^\dagger(\vec{r})$ on the vacuum:

$$\hat{\Psi}^\dagger(\vec{r})|0\rangle = \sum_\alpha \phi_\alpha^T(\vec{r})c_\alpha^\dagger|0\rangle = \sum_\alpha \phi_\alpha^T(\vec{r})|0, \dots, n_\alpha = 1, 0, \dots\rangle$$

I can also write $|0, \dots, n_\alpha = 1, 0, \dots\rangle = |\alpha\rangle$, since this is the state with a single particle in state α , and so it follows that:

$$\langle \vec{r}' | \hat{\Psi}^\dagger(\vec{r}) | 0 \rangle = \sum_\alpha \phi_\alpha^T(\vec{r}) \phi_\alpha(\vec{r}') = \delta(\vec{r} - \vec{r}')$$

since the basis is complete. This is possible for any \vec{r}, \vec{r}' only iff:

$$\hat{\Psi}^\dagger(\vec{r})|0\rangle = |\vec{r}\rangle$$

i.e. this operator creates a particle at position \vec{r} (with any spin! That is why these operators are spinors; we'll specialize to a specific spin projection below). It can be similarly shown that this is true in any state (not only vacuum). Similarly, it can be shown that the operator $\hat{\Psi}(\vec{r})$ destroys (removes) a particle from position \vec{r} . So these operators are simply the creation and annihilation operators associated with the basis $|\vec{r}\rangle$; we give them special names and notation just because we're biased in thinking that "space" is somehow more special than other representations.

Using again the completeness and orthonormality of the basis α , we can invert Eqs. (19) to find:

$$c_\alpha = \int d\vec{r} \phi_\alpha^T(\vec{r}) \hat{\Psi}(\vec{r}); \quad c_\alpha^\dagger = \int d\vec{r} \hat{\Psi}^\dagger(\vec{r}) \phi_\alpha(\vec{r}) \quad (21)$$

and substituting this in Eqs. (8), (10) we find the representations:

$$\hat{A} = \sum_{i=1}^N \mathcal{A}_i \rightarrow \int d\vec{r} \hat{\Psi}^\dagger(\vec{r}) \mathcal{A}_{\vec{r}} \hat{\Psi}(\vec{r}) \quad (22)$$

and

$$\vec{B} = \frac{1}{2} \sum_{i \neq j} \mathcal{B}_{ij} \rightarrow \frac{1}{2} \int d\vec{r} \int d\vec{r}' \hat{\Psi}^\dagger(\vec{r}) \hat{\Psi}^\dagger(\vec{r}') \mathcal{B}_{\vec{r}, \vec{r}'} \hat{\Psi}(\vec{r}') \hat{\Psi}(\vec{r}) \quad (23)$$

(note the order of operators!!!), where I used the simplified notation $\mathcal{A}_{\vec{r}} = \langle \vec{r} | \mathcal{A} | \vec{r} \rangle$ – this is a $2S + 1 \times 2S + 1$ matrix, because of the spin – and similarly for B (which is a direct product of two $2S + 1 \times 2S + 1$ matrices). (Don't worry, we'll soon introduce friendlier notation, so we won't need to deal with these matrices. But let's be formal just for a bit more.)

It follows that:

(a) $\hat{T} = \int d\vec{r} \hat{\Psi}^\dagger(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 \right) \hat{\Psi}(\vec{r});$

(b) density operator $\hat{n}(\vec{r}) = \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i) \rightarrow \int d\vec{r}' \hat{\Psi}^\dagger(\vec{r}') \delta(\vec{r} - \vec{r}') \hat{\Psi}(\vec{r}') = \hat{\Psi}^\dagger(\vec{r}) \hat{\Psi}(\vec{r})$

(c) total number of particles $\hat{N} = \sum_{i=1}^N 1 \rightarrow \int d\vec{r} \hat{\Psi}^\dagger(\vec{r}) \hat{\Psi}(\vec{r}) = \int d\vec{r} \hat{n}(\vec{r})$

(d) Coulomb interaction: $\hat{V} = \frac{1}{2} \int d\vec{r} \int d\vec{r}' \hat{\Psi}^\dagger(\vec{r}) \hat{\Psi}^\dagger(\vec{r}') u(\vec{r} - \vec{r}') \hat{\Psi}(\vec{r}') \hat{\Psi}(\vec{r}),$ etc.

In particular, the Hamiltonian for particles in an external field and also interacting with one another becomes:

$$\mathcal{H} = \int d\vec{r} \hat{\Psi}^\dagger(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 + u_{ext}(\vec{r}) \right) \hat{\Psi}(\vec{r}) + \frac{1}{2} \int d\vec{r} \int d\vec{r}' \hat{\Psi}^\dagger(\vec{r}) \hat{\Psi}^\dagger(\vec{r}') u(\vec{r} - \vec{r}') \hat{\Psi}(\vec{r}') \hat{\Psi}(\vec{r}) \quad (24)$$

looks very appealing, and is again independent on the number of particles that may be in the system, or whether it is fixed or not. As I said, this makes it very easy to work with states in which the number of particles is not fixed, such as is the case in grand-canonical ensembles, for instance.

When you will learn about Green's functions (in phys503) you will derive the equation of motion for the field operator $\hat{\Psi}(\vec{r})$ and see that it is very similar in form to Schrödinger's equation. In the absence of particle-particle interactions, it turns out that $i\hbar \frac{d}{dt} \hat{\Psi}(\vec{r}, t) = \left(-\frac{\hbar^2}{2m} \nabla^2 + u_{ext}(\vec{r}) \right) \hat{\Psi}(\vec{r}, t)$. Interactions add a second term. The reason I'm mentioning this now, is that it gives you a sense of why this is called the second quantization. In the first quantization (going from classical to quantum mechanics) we quantize the operators, and use a wavefunction $\phi(\vec{r})$ to characterize the state of the system. This wavefunction $\phi(\vec{r})$ can be regarded as a "classical field" of its own. In the second quantization, it is as if we now also quantize the wavefunction into the field operator $\hat{\Psi}(\vec{r})$.

Now, let us simplify the notation and remove the need of working with matrices if the spin $S > 0$. In all cases we can choose the basis $\alpha = (\tilde{\alpha}, \sigma)$, i.e. one of the quantum numbers is the spin, and then $\tilde{\alpha}$ contains all other quantum numbers needed to describe the translational sector of the Hilbert space (e.g., the momentum of the particle, or hydrogen orbitals, or whatever).

We can then define *spin-components of the field operators*:

$$\hat{\Psi}_\sigma(\vec{r}) = \sum_{\tilde{\alpha}} \phi_{\tilde{\alpha}}(\vec{r}) c_{\tilde{\alpha}, \sigma}; \quad \hat{\Psi}_\sigma^\dagger(\vec{r}) = \sum_{\tilde{\alpha}} \phi_{\tilde{\alpha}}^*(\vec{r}) c_{\tilde{\alpha}, \sigma}^\dagger \quad (25)$$

If $\tilde{\alpha} = \vec{k}$, then $\phi_{\tilde{\alpha}}(\vec{r}) = e^{i\vec{r}\vec{k}}/\sqrt{V}$ and is a simple number, not a $2S + 1$ -dim spinor. So the spin-components $\hat{\Psi}_\sigma(\vec{r}), \hat{\Psi}_\sigma^\dagger(\vec{r})$ are simple operators, not vectors of operators. Their meaning is that they annihilate or create a particle *with spin* σ at position \vec{r} .

It is now straightforward to verify (by repeating previous calculations) that:

$$[\hat{\Psi}_\sigma(\vec{r}), \hat{\Psi}_{\sigma'}^\dagger(\vec{r}')]_\xi = \delta_{\sigma, \sigma'} \delta(\vec{r} - \vec{r}'); \quad [\hat{\Psi}_\sigma(\vec{r}), \hat{\Psi}_{\sigma'}(\vec{r}')]_\xi = [\hat{\Psi}_\sigma^\dagger(\vec{r}), \hat{\Psi}_{\sigma'}^\dagger(\vec{r}')]_\xi = 0 \quad (26)$$

and to show that:

$$\hat{A} = \sum_{i=1}^N \mathcal{A}_i \rightarrow \sum_{\sigma, \sigma'} \int d\vec{r} \hat{\Psi}_\sigma^\dagger(\vec{r}) \langle \sigma | \mathcal{A}_i | \sigma' \rangle \hat{\Psi}_{\sigma'}(\vec{r}) \quad (27)$$

and

$$\vec{B} = \frac{1}{2} \sum_{i \neq j} \mathcal{B}_{ij} \rightarrow \frac{1}{2} \sum_{\substack{\sigma_1, \sigma_2 \\ \sigma_3, \sigma_4}} \int d\vec{r} \int d\vec{r}' \hat{\Psi}_{\sigma_1}^\dagger(\vec{r}) \hat{\Psi}_{\sigma_2}^\dagger(\vec{r}') \langle \sigma_1, \sigma_2 | \mathcal{B}_{\vec{r}, \vec{r}'} | \sigma_3, \sigma_4 \rangle \hat{\Psi}_{\sigma_4}(\vec{r}') \hat{\Psi}_{\sigma_3}(\vec{r}) \quad (28)$$

where $\langle \sigma | \mathcal{A}_{\vec{r}} | \sigma' \rangle = \langle \vec{r}, \sigma | \mathcal{A} | \vec{r}, \sigma' \rangle$ and $\langle \sigma_1, \sigma_2 | \mathcal{B}_{\vec{r}, \vec{r}'} | \sigma_3, \sigma_4 \rangle = \langle \vec{r}, \sigma_1; \vec{r}', \sigma_2 | \mathcal{B} | \vec{r}, \sigma_3; \vec{r}', \sigma_4 \rangle$. In particular, if the operators are spin independent, then these matrix elements are proportional to $\delta_{\sigma, \sigma'}$, respectively $\delta_{\sigma_1, \sigma_3} \delta_{\sigma_2, \sigma_4}$. As a result, a typical Hamiltonian becomes:

$$\mathcal{H} = \sum_{\sigma} \int d\vec{r} \hat{\Psi}_{\sigma}^{\dagger}(\vec{r}) \left(-\frac{\hbar^2}{2m} \nabla^2 + u_{ext}(\vec{r}) \right) \hat{\Psi}_{\sigma}(\vec{r}) + \frac{1}{2} \sum_{\sigma, \sigma'} \int d\vec{r} \int d\vec{r}' \hat{\Psi}_{\sigma}^{\dagger}(\vec{r}) \hat{\Psi}_{\sigma'}^{\dagger}(\vec{r}') u(\vec{r} - \vec{r}') \hat{\Psi}_{\sigma'}(\vec{r}') \hat{\Psi}_{\sigma}(\vec{r}) \quad (29)$$