# 2 Mean-Field Theory: Hartree-Fock and BCS

Erik Koch Institute for Advanced Simulation Forschungszentrum Jülich

# Contents

1	Man	y-electron states	2								
2	Second quantization										
	2.1	Creation and annihilation operators	5								
	2.2	Representation of Slater determinants	6								
	2.3	Representation of <i>n</i> -body operators	7								
	2.4	Transforming the orbital basis	11								
3	Variational methods										
	3.1	Non-interacting electrons	14								
	3.2	Hartree-Fock theory	15								
	3.3	BCS theory	24								
4	Con	clusion	27								
A	Appendix										
	A.1	Non-orthonormal basis	28								
	A.2	Pauli matrices	29								
	A.3	Some useful commutation relations	31								

E. Pavarini and E. Koch (eds.) Correlations and Phase Transitions Modeling and Simulation Vol. 14 Forschungszentrum Jülich, 2024, ISBN 978-3-95806-751-6 http://www.cond-mat.de/events/correl24

## **1** Many-electron states

One of the great surprises of quantum mechanics is the existence of indistinguishable objects. Classically this is not possible: objects can always be distinguished at least by their position in space, meaning that indistinguishable objects must be identical. This is Leibniz' Principle of the Identity of Indiscernibles [1]. For quantum objects, however, the uncertainty principle makes the distinction of particles by their position impossible. This allows for the existence of elementary particles. They form the basic units of all matter. So, quite remarkably, all the different objects we know are made of indistinguishable building blocks.

In the formalism of quantum mechanics, indistinguishability means that no observable lets us distinguish one of these particles from the other. This means that every observable for, e.g., electrons, must treat each electron in the same way. Thus, in principle, observables must act on all electrons in the universe. In practice we can, of course, distinguish electrons localized on the moon from those in our lab to an excellent approximation. Thus, for all practical purposes, we can restrict our description to the electrons in the system under consideration, assuming that the differential overlap with all other electrons vanishes. Any observable  $M(x_1, \ldots, x_N)$  for the N electrons in our system must then be symmetric under permutations of the variables  $x_i$ .

The consequences are straightforward: An observable M(x) acting on a single-particle degree of freedom x must act on all indistinguishable particles in the same way, i.e.,  $\sum_{i} M(x_i)$ . Likewise, a two-body observable M(x, x') must act on all pairs in the same way,  $\sum_{i,j} M(x_i, x_j)$ with M(x, x') = M(x', x). We can thus write any observable in the form

$$M(\mathbf{x}) = M^{(0)} + \sum_{i} M^{(1)}(x_i) + \frac{1}{2!} \sum_{i \neq j} M^{(2)}(x_i, x_j) + \frac{1}{3!} \sum_{i \neq j \neq k} M^{(3)}(x_i, x_j, x_k) + \cdots$$
(1)

$$= M^{(0)} + \sum_{i} M^{(1)}(x_i) + \sum_{i < j} M^{(2)}(x_i, x_j) + \sum_{i < j < k} M^{(3)}(x_i, x_j, x_k) + \cdots, \quad (2)$$

where the summations can be restricted since the operators must be symmetric in their arguments, while for two or more identical coordinates the operator is really one of lower order:  $M^{(2)}(x_i, x_i)$ , e.g., only acts on a single coordinate and should be included in  $M^{(1)}$ .

For the many-body wavefunctions  $\Psi(x_1, x_2, \dots)$  the situation is slightly more complex. Since the probability density  $|\Psi(x_1, x_2, \dots)|^2$  is an observable, the wavefunction should transform as one-dimensional (irreducible) representations of the permutation group. Which irreducible representation applies to a given type of elementary particle is determined by the spin-statistics theorem [2,3]: The wavefunctions of particles with integer spin are symmetric, those of particles with half-integer spin change sign when two arguments are exchanged. From an arbitrary *N*particle wavefunction we thus obtain a many-electron wavefunction by antisymmetrizing

$$\mathcal{A}\Psi(x_1,\ldots,x_N) := \frac{1}{\sqrt{N!}} \sum_P (-1)^P \Psi\left(x_{p(1)},\ldots,x_{p(N)}\right), \qquad (3)$$

where  $(-1)^P$  is the parity of the permutation P that maps  $n \to p(n)$ . Since there are N! different permutations, this can easily become an extremely costly operation. Remarkably, a product of

N single-electron states  $\varphi_{\alpha}$  can be antisymmetrized much more efficiently (in  $\mathcal{O}(N^3)$  steps) by writing it in the form of a determinant

$$\Phi_{\alpha_1,\dots,\alpha_N}(x_1,\dots,x_N) := \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix}.$$
(4)

For N=1 the Slater determinant is simply the one-electron orbital  $\Phi_{\alpha}(x) = \varphi_{\alpha}(x)$  while for N=2 we get the familiar expression  $\Phi_{\alpha,\alpha'}(x,x') = (\varphi_{\alpha}(x)\varphi_{\alpha'}(x') - \varphi_{\alpha'}(x)\varphi_{\alpha}(x'))/\sqrt{2}$  for the two-electron Slater determinant.

Slater determinants are important because they can be used to build a basis of the many-electron Hilbert space. To see how, we consider a complete set of orthonormal single-electron states

$$\sum_{n} \overline{\varphi_n(x)} \,\varphi_n(x') = \delta(x - x') \text{ (complete)} \quad \int dx \,\overline{\varphi_n(x)} \,\varphi_m(x) = \delta_{n,m} \text{ (orthonormal)}. \tag{5}$$

To expand an arbitrary N-particle function  $a(x_1, \ldots, x_N)$ , we start by considering it as a function of  $x_1$  with  $x_2, \ldots, x_N$  kept fixed. We can then expand it in the complete set  $\{\varphi_n\}$  as

$$a(x_1, \dots, x_N) = \sum_{n_1} a_{n_1}(x_2, \dots, x_N) \varphi_{n_1}(x_1)$$

with expansion coefficients that depend on the remaining coordinates

$$a_{n_1}(x_2,\ldots,x_N) = \int dx_1 \ \overline{\varphi_{n_1}(x_1)} \ a(x_1,x_2,\ldots,x_N).$$

These, in turn, can be expanded as a functions of  $x_2$ 

$$a_{n_1}(x_2,\ldots,x_N) = \sum_{n_2} a_{n_1,n_2}(x_3,\ldots,x_N) \varphi_{n_2}(x_2).$$

Repeating this, we obtain the expansion of a in product states

$$a(x_1,\ldots,x_N) = \sum_{n_1,\ldots,n_N} a_{n_1,\ldots,n_N} \varphi_{n_1}(x_1) \cdots \varphi_{n_N}(x_N)$$

with

$$a_{n_1,\dots,n_N} = \int dx_1 \cdots \int dx_N \,\overline{\varphi_{n_1}(x_1)} \cdots \overline{\varphi_{n_N}(x_N)} \, a(x_1,\dots,x_N).$$

For an antisymmetric N-particle function  $\Psi$  the expansion coefficients are thus antisymmetric under permutation of the indices:  $a_{n_{p(1)},...,n_{p(N)}} = (-1)^{P} a_{n_{1},...,n_{N}}$ . Fixing some particular order of the indices, e.g.,  $n_{1} < n_{2} < ... < n_{N}$ , we thus get an expansion in Slater determinants

$$\Psi(x_1, \dots, x_N) = \sum_{n_1 < \dots < n_N} a_{n_1, \dots, n_N} \sqrt{N!} \, \Phi_{n_1, \dots, n_N}(x_1, \dots, x_N).$$

Since we can write any antisymmetric function as such a configuration-interaction expansion, the set of Slater determinants

$$\left\{ \Phi_{n_1,\dots,n_N}(x_1,\dots,x_N) \mid n_1 < n_2 < \dots < n_N \right\}$$
 (6)

forms a basis of the N-electron Hilbert space. Since the overlap of two Slater determinants

$$\int d\boldsymbol{x} \,\overline{\Phi_{\alpha_1,\dots,\alpha_N}(\boldsymbol{x})} \,\Phi_{\beta_1,\dots,\beta_N}(\boldsymbol{x}) = \frac{1}{N!} \sum_{P,P'} (-1)^{P+P'} \prod_n \int dx_n \,\overline{\varphi_{\alpha_{p(n)}}(x_n)} \,\varphi_{\alpha_{p'(n)}}(x_n)$$
$$= \begin{vmatrix} \langle \varphi_{\alpha_1} | \varphi_{\beta_1} \rangle & \cdots & \langle \varphi_{\alpha_1} | \varphi_{\beta_N} \rangle \\ \vdots & \ddots & \vdots \\ \langle \varphi_{\alpha_N} | \varphi_{\beta_1} \rangle & \cdots & \langle \varphi_{\alpha_N} | \varphi_{\beta_N} \rangle \end{vmatrix}$$
(7)

is the determinant of the overlap of the constituent orbitals, the Slater determinants (6) form a complete orthonormal basis of the N-electron Hilbert space when the orbitals  $\varphi_n(x)$  are a complete orthonormal basis of the one-electron Hilbert space.

While we use a set of N one-electron orbitals  $\varphi_n(x)$  to define an N-electron Slater determinant  $\Phi_{\alpha_1,\ldots,\alpha_N}(x)$ , this representation is not unique: Any unitary transformation among the N occupied orbitals will not change the determinant. Thus, strictly, a Slater determinant is not determined by the set of indices we usually give, but, up to a phase, by the N-dimensional subspace spanned by the orbitals  $\varphi_1, \ldots, \varphi_N$  in the single-electron Hilbert space. The projector to this space is the one-body density matrix

$$\Gamma^{(1)}(x,x') = N \int dx_2 \cdots dx_N \,\overline{\Phi(x,x_2,\ldots,x_N)} \,\Phi(x',x_2,\ldots,x_N) \,. \tag{8}$$

To see this, we expand the Slater determinant along its first row

$$\Phi_{\alpha_1 \cdots \alpha_N}(x_1, \dots, x_N) = \frac{1}{\sqrt{N}} \sum_{n=1}^N (-1)^{1+n} \varphi_{\alpha_n}(x_1) \Phi_{\alpha_{i\neq n}}(x_2, \dots, x_N), \qquad (9)$$

where  $\Phi_{\alpha_{i\neq n}}(x_2, \ldots, x_N)$  is the determinant with the first row and the *n*-th column removed, which can be written as N-1-electron Slater determinants with orbital  $\alpha_n$  removed. Inserting this into (8) we find

$$\Gamma_{\Phi}^{(1)}(x,x') = \sum_{n=1}^{N} \overline{\varphi_{\alpha_n}(x)} \,\varphi_{\alpha_n}(x') \,, \tag{10}$$

which is the expansion of the one-body density matrix in eigenfunctions (natural orbitals), with eigenvalues (natural occupation numbers) either one or zero. Any many-electron wavefunction  $\Psi(\boldsymbol{x})$  with the same one-body density matrix  $\Gamma_{\Phi}^{(1)}$  equals the Slater determinant  $\Phi(\boldsymbol{x})$  up to a phase, i.e.,  $|\langle \Psi | \Phi \rangle| = 1$ .

We can generalize this procedure and calculate higher order density matrices by introducing the generalized Laplace expansion

$$\Phi_{\alpha_1 \cdots \alpha_N}(\mathbf{x}) = \frac{1}{\sqrt{\binom{N}{p}}} \sum_{n_1 < \cdots < n_p} (-1)^{1 + \sum_i n_i} \Phi_{\alpha_{n_1} \cdots \alpha_{n_p}}(x_1, \dots, x_p) \Phi_{\alpha_{i \notin \{n_1, \dots, n_p\}}}(x_{p+1}, \dots, x_N),$$

which is obtained by writing the permutation of all N indices as a permutation of N-p indices and the remaining p indices separately, summing over all distinct sets of p indices. This allows us to evaluate arbitrary matrix elements and higher order density matrices [4]. But as can be seen from the above expansion, the expressions very quickly get quite cumbersome. Fortunately there is a representation that is much better suited to handling antisymmetric wavefunctions. It is called second quantization.

## 2 Second quantization

Second quantization is the generalization of the Dirac notation to many-electron states [5, 6]. The idea is to separate the wavefunction into coordinate and state,  $\varphi_{\alpha}(x) = \langle x | \alpha \rangle$ , and absorb the coordinate in the operators, so that expectation values are rewritten as

$$\int dx \,\overline{\varphi_{\alpha}(x)} M(x) \varphi_{\beta}(x') = \langle \alpha | \int dx \, |x\rangle M(x) \langle x | \beta \rangle =: \langle \alpha | \hat{M} | \beta \rangle.$$

For many-electron systems the coordinates are represented by the field-operators  $\hat{\Psi}(x)$ . To implement antisymmetry, they need to anticommute  $\hat{\Psi}(x)\hat{\Psi}(x') = -\hat{\Psi}(x')\hat{\Psi}(x)$ . They can be thought of as removing an electron with coordinates  $x=(r,\sigma)$  from the system. Applying them on a state with no electrons thus gives zero. Their conjugate operators insert an electron to the system. The entire formalism is defined by a compact set of properties of the vacuum state  $|0\rangle$  and the anticommutation relations of the field operators:

$$\hat{\Psi}(x)|0\rangle = 0 \qquad \{\hat{\Psi}(x), \hat{\Psi}(x')\} = 0 = \{\hat{\Psi}^{\dagger}(x), \hat{\Psi}^{\dagger}(x')\} 
\langle 0|0\rangle = 1 \qquad \{\hat{\Psi}(x), \hat{\Psi}^{\dagger}(x')\} = \delta(x - x')$$
(11)

where the curly braces denote the anti-commutator  $\{A, B\} := AB + BA$ . As a direct consequence we obtain the Pauli principle:  $\hat{\Psi}^{\dagger}(x)\hat{\Psi}^{\dagger}(x) = 0$ .

#### 2.1 Creation and annihilation operators

For a single-electron state  $\varphi_{\alpha}(x)$  we can define an operator

$$c_{\alpha}^{\dagger} := \int dx \,\varphi_{\alpha}(x) \,\hat{\Psi}^{\dagger}(x), \tag{12}$$

that inserts an electron at position x weighted with the amplitude  $\varphi_{\alpha}(x)$  — hence called creation operator for state  $|\alpha\rangle$ . Note that the creation operators transform in the same way as the singleelectron states they represent, not like operators in first quantization,

$$|\tilde{\alpha}_i\rangle = \sum_{\mu} |\alpha_{\mu}\rangle U_{\mu i} \qquad \rightsquigarrow \quad \tilde{c}^{\dagger}_{\tilde{\alpha}_i}|0\rangle = \sum_{\mu} c^{\dagger}_{\alpha_{\mu}}|0\rangle U_{\mu i} = \left(\sum_{\mu} c^{\dagger}_{\alpha_{\mu}}U_{\mu i}\right)|0\rangle.$$
(13)

The anticommutator with the field annihilator just gives back the single-electron wavefunction

$$\left\{\hat{\Psi}(x),\,c^{\dagger}_{\alpha}\right\} = \int dx'\,\varphi_{\alpha}(x')\left\{\hat{\Psi}(x),\,\hat{\Psi}^{\dagger}(x')\right\} = \varphi_{\alpha}(x)\,. \tag{14}$$

Thus, together with their adjoints, the annihilation operators,

$$c_{\alpha} := \int dx \,\overline{\varphi_{\alpha}(x)} \,\hat{\Psi}(x),\tag{15}$$

they form the algebra

$$c_{\alpha}|0\rangle = 0 \qquad \left\{c_{\alpha}, c_{\beta}\right\} = 0 = \left\{c_{\alpha}^{\dagger}, c_{\beta}^{\dagger}\right\}$$

$$\langle 0|0\rangle = 1 \qquad \left\{c_{\alpha}, c_{\beta}^{\dagger}\right\} = \langle \alpha|\beta\rangle$$
(16)

The notation of second quantization has two principal advantages: (i) it is impossible to write unphysical states which are not antisymmetric or operators which are not acting on all electrons in the same way, and (ii) operators are independent of particle number, so that it becomes straightforward to work in Fock space. This simplifies many-body calculations significantly.

#### 2.2 Representation of Slater determinants

We have now all the tools in place to write the Slater determinant (4) in second quantization, using the creation operators to specify the occupied orbitals and the field operators to give the coordinates for the real-space representation:

$$\Phi_{\alpha_1\alpha_2\dots\alpha_N}(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \left\langle 0 \left| \hat{\Psi}(x_1) \hat{\Psi}(x_2) \dots \hat{\Psi}(x_N) \right| c_{\alpha_N}^{\dagger} \dots c_{\alpha_2}^{\dagger} c_{\alpha_1}^{\dagger} \right| 0 \right\rangle.$$
(17)

Note how writing the Slater determinant as an expectation value of annihilation and creation operators nicely separates the coordinates on the left from the orbitals on the right. This is just the desired generalization of the Dirac notation  $\varphi(x) = \langle x | \varphi \rangle$ .

Not surprisingly, the proof of (17) is by induction. As a warm-up we consider the case of a single-electron wavefunction (N = 1). Using the anticommutation relation (14), we see that

$$\left\langle 0 \left| \hat{\Psi}(x_1) c_{\alpha_1}^{\dagger} \right| 0 \right\rangle = \left\langle 0 \left| \varphi_{\alpha_1}(x_1) - c_{\alpha_1}^{\dagger} \hat{\Psi}(x_1) \right| 0 \right\rangle = \varphi_{\alpha_1}(x_1).$$
(18)

For the two-electron state N = 2, we anticommute  $\hat{\Psi}(x_2)$  in two steps to the right

$$\begin{aligned} \left\langle 0 \right| \hat{\Psi}(x_{1}) \hat{\Psi}(x_{2}) c_{\alpha_{2}}^{\dagger} c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle &= \left\langle 0 \right| \hat{\Psi}(x_{1}) \left( \varphi_{\alpha_{2}}(x_{2}) - c_{\alpha_{2}}^{\dagger} \hat{\Psi}(x_{2}) \right) c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle \\ &= \left\langle 0 \right| \hat{\Psi}(x_{1}) c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle \varphi_{\alpha_{2}}(x_{2}) - \left\langle 0 \right| \hat{\Psi}(x_{1}) c_{\alpha_{2}}^{\dagger} \hat{\Psi}(x_{2}) c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle \\ &= \varphi_{\alpha_{1}}(x_{1}) \varphi_{\alpha_{2}}(x_{2}) - \varphi_{\alpha_{2}}(x_{1}) \varphi_{\alpha_{1}}(x_{2}). \end{aligned}$$
(19)

We see how anticommuting automatically produces the appropriate signs for the antisymmetric wavefunction. Dividing by  $\sqrt{2}$ , we obtain the desired two-electron Slater determinant. The general case of an N-electron state works just the same. Anti-commuting  $\hat{\Psi}(x_N)$  all the way to the right produces N-1 terms with alternating sign

$$\begin{aligned} \left\langle 0 \right| \hat{\Psi}(x_{1}) \cdots \hat{\Psi}(x_{N-1}) \hat{\Psi}(x_{N}) c_{\alpha_{N}}^{\dagger} c_{\alpha_{N-1}}^{\dagger} \cdots c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle &= \\ &+ \left\langle 0 \right| \hat{\Psi}(x_{1}) \cdots \hat{\Psi}(x_{N-1}) c_{\alpha_{N-1}}^{\dagger} \cdots c_{\alpha_{1}}^{\dagger} \left| 0 \right\rangle \quad \varphi_{\alpha_{N}} \left( x_{N} \right) \\ &- \left\langle 0 \right| \hat{\Psi}(x_{1}) \cdots \hat{\Psi}(x_{N-1}) \prod_{n \neq N-1} c_{\alpha_{n}}^{\dagger} \left| 0 \right\rangle \quad \varphi_{\alpha_{N-1}}(x_{N}) \\ &\vdots \\ &(-1)^{N-1} \left\langle 0 \right| \hat{\Psi}(x_{1}) \cdots \hat{\Psi}(x_{N-1}) c_{\alpha_{N}}^{\dagger} \cdots c_{\alpha_{2}}^{\dagger} \left| 0 \right\rangle \quad \varphi_{\alpha_{1}} \left( x_{N} \right). \end{aligned}$$

Using (17) for the N-1-electron states, this is nothing but the Laplace expansion of

$$D = \begin{vmatrix} \varphi_{\alpha_1}(x_1) & \varphi_{\alpha_2}(x_1) & \cdots & \varphi_{\alpha_N}(x_1) \\ \varphi_{\alpha_1}(x_2) & \varphi_{\alpha_2}(x_2) & \cdots & \varphi_{\alpha_N}(x_2) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_{\alpha_1}(x_N) & \varphi_{\alpha_2}(x_N) & \cdots & \varphi_{\alpha_N}(x_N) \end{vmatrix}$$

along the Nth row. Dividing by  $\sqrt{N!}$  we see that we have shown (17) for N-electron states, completing the proof by induction.

Given this representation of Slater determinants it is easy to eliminate the coordinates so we can work with N-electron states rather than N-electron wavefunctions—just as in Dirac notation. In particular we can rewrite the basis of Slater determinants (6) into a basis of product states

$$\left\{ c_{n_N}^{\dagger} \cdots c_{n_1}^{\dagger} | 0 \right\rangle \mid n_1 < \cdots < n_N \right\}, \tag{20}$$

which allows us to express any N-electron state as

$$|\Psi\rangle = \sum_{n_1 < \dots < n_N} a_{n_1,\dots,n_N} c^{\dagger}_{n_N} \cdots c^{\dagger}_{n_1} |0\rangle.$$
(21)

#### **2.3** Representation of *n*-body operators

To work with N-electron states rather than Slater determinants, we also have to rewrite the N-electron operators M(x) appropriately. This is easily done by incorporating the coordinates that we have separated from the Slater determinants into the operators such that the expectation values remain unchanged. This is, again, analogous to the Dirac formalism:

$$\int dx \,\overline{\varphi_n(x)} \, M(x) \,\varphi_m(x) = \langle \varphi_n | \underbrace{\int dx \, |x\rangle M(x) \langle x|}_{=:\hat{M}} \varphi_m \rangle = \langle \varphi_n | \hat{M} | \varphi_m \rangle. \tag{22}$$

For N-electron Slater determinants it becomes

$$\int dx_1 \cdots dx_N \,\overline{\Phi_{\beta_1 \cdots \beta_N}(x_1, \cdots, x_N)} \, M(x_1, \dots, x_N) \, \Phi_{\alpha_1 \cdots \alpha_N}(x_1, \cdots, x_N)$$
$$= \int dx_1 \cdots dx_N \langle 0 | c_{\beta_1} \cdots c_{\beta_N} \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_1) | 0 \rangle M(x_1, \dots, x_N) \langle 0 | \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) c_{\alpha_N}^{\dagger} \cdots c_{\alpha_1}^{\dagger} | 0 \rangle$$
$$= \langle 0 | c_{\beta_1} \cdots c_{\beta_N} \, \hat{M} \, c_{\alpha_N}^{\dagger} \cdots c_{\alpha_1}^{\dagger} | 0 \rangle$$

with the representation of the n-body operator in terms of field operators

$$\hat{M} := \frac{1}{N!} \int dx_1 \cdots x_N \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_1) M(x_1, \cdots, x_N) \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) .$$
(23)

Note that this particular form of the operator is only valid when applied to N-electron states, since we have used that the N annihilation operators bring us to the zero-electron space, where  $|0\rangle\langle 0| = 1$ . Keeping this in mind, we can work entirely in terms of our algebra (11).

To see what (23) means, we look, in turn, at the different *n*-body parts of M(x), (2):

$$M(\mathbf{x}) = M^{(0)} + \sum_{i} M^{(1)}(x_i) + \sum_{i < j} M^{(2)}(x_i, x_j) + \sum_{i < j < k} M^{(3)}(x_i, x_j, x_k) + \cdots$$
(24)

We start with the simplest case, the zero-body operator, which, just a number  $M^{(0)}$ . Operating on an N-electron wavefunction, it gives

$$\hat{M}^{(0)} = \frac{M^{(0)}}{N!} \int dx_1 dx_2 \cdots x_N \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_2) \hat{\Psi}^{\dagger}(x_1) \hat{\Psi}(x_1) \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N)$$

$$= \frac{M^{(0)}}{N!} \int dx_2 \cdots x_N \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_2) \qquad \hat{N} \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N)$$

$$= \frac{M^{(0)}}{N!} \int dx_2 \cdots x_N \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_2) \qquad 1 \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N)$$

$$\vdots$$

$$= \frac{M^{(0)}}{N!} 1 \cdot 2 \cdots N = M^{(0)}, \qquad (25)$$

where we have used that the operator

$$\int dx \,\hat{\Psi}^{\dagger}(x)\hat{\Psi}(x) = \hat{N}$$

counts the creation/annihilation operators: Applied to the vacuum state it gives  $\hat{N} |0\rangle = 0$ , while its commutator with any creation operator produces that operator

$$[\hat{N}, c_n^{\dagger}] = \int dx \, [\hat{\Psi}^{\dagger}(x)\hat{\Psi}(x), c_n^{\dagger}] = \int dx \, \hat{\Psi}^{\dagger}(x) \, \{\hat{\Psi}(x), c_n^{\dagger}\} = \int dx \, \hat{\Psi}^{\dagger}(x) \, \varphi_n(x) = c_n^{\dagger}.$$
(26)

where we have used the simple relation  $[AB, C] = A\{B, C\} - \{A, C\}B$ . Commuting with an annihilator we pick up a minus sign  $[\hat{N}, \hat{\Psi}(x)] = -\hat{\Psi}(x)$ , i.e.,  $\hat{N}\hat{\Psi}(x) = \hat{\Psi}(x)(\hat{N}-1)$ . Thus, commuting  $\hat{N}$  through a general product state, we obtain for each creation operator that we encounter a copy of the state, while for each annihilator we obtain minus that state, giving in total the original state times the difference in the number of creation and annihilation operators. Remarkably, while we started from an operator acting on *N*-electron states, the resulting operator in second quantized form is independent of the number of electrons. We will see that this is an important general feature of operators in second quantization which makes working in Fock spaces amazingly simple.

We note that (25) just means that the overlap of two Slater determinants (7) is equal to that of the corresponding product states

$$\int d\boldsymbol{x} \,\overline{\Phi_{\alpha_1,\dots,\alpha_N}(\boldsymbol{x})} \,\Phi_{\beta_1,\dots,\beta_N}(\boldsymbol{x}) = \left\langle 0 \right| c_{\alpha_1} \cdots c_{\alpha_N} \,c_{\beta_N}^{\dagger} \cdots c_{\beta_1}^{\dagger} \left| 0 \right\rangle. \tag{27}$$

#### 2.3.1 One-body operators

Next we consider one-body operators  $\sum_{j} M^{(1)}(x_j)$ 

$$\hat{M}^{(1)} = \frac{1}{N!} \int dx_1 \cdots dx_N \,\hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_1) \sum_j M^{(1)}(x_j) \,\hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N)$$

$$= \frac{1}{N!} \sum_j \int dx_j \,\hat{\Psi}^{\dagger}(x_j) \,M^{(1)}(x_j) \,(N-1)! \,\hat{\Psi}(x_j)$$

$$= \frac{1}{N!} \sum_j \int dx_j \,\hat{\Psi}^{\dagger}(x_j) \,M^{(1)}(x_j) \,\hat{\Psi}(x_j)$$

$$= \int dx \,\,\hat{\Psi}^{\dagger}(x) \,\,M^{(1)}(x) \,\,\hat{\Psi}(x)$$
(28)

Here we have first anticommuted  $\hat{\Psi}^{\dagger}(x_j)$  all the way to the left and  $\hat{\Psi}(x_j)$  to the right. Since these take the same numbers of anticommutations, there is no sign involved. In between these field operators we are left with a zero-body operator for N-1 electrons, producing, when  $\hat{M}^{(1)}$ acts on an N-electron state, a factor of (N-1)!. Again we notice that we obtain an operator that no longer depends on the number of electrons, i.e., that is valid in the entire Fock space. Expanding the field-operators in a complete orthonormal set  $\hat{\Psi}(x) = \sum_{n} \varphi_n(x) c_n$  gives

$$\hat{M}^{(1)} = \sum_{n,m} \int dx \,\overline{\varphi_n(x)} \, M(x) \,\varphi_m(x) \, c_n^{\dagger} c_m = \sum_{n,m} \langle \varphi_n | M^{(1)} | \varphi_m \rangle \, c_n^{\dagger} c_m = \sum_{n,m} c_n^{\dagger} \, M_{nm}^{(1)} \, c_m. \tag{29}$$

The matrix elements  $M_{nm}^{(1)} = \langle \varphi_n | M^{(1)} | \varphi_m \rangle$  transform like a single-electron matrix  $M^{(1)}$ : From (13) and writing the annihilation operators as a column vector c we see that

$$\hat{M}^{(1)} = \boldsymbol{c}^{\dagger} \boldsymbol{M}^{(1)} \boldsymbol{c} = \boldsymbol{c}^{\dagger} \boldsymbol{U}^{\dagger} \boldsymbol{U} \boldsymbol{M}^{(1)} \boldsymbol{U}^{\dagger} \boldsymbol{U} \boldsymbol{c} = \tilde{\boldsymbol{c}}^{\dagger} \tilde{\boldsymbol{M}}^{(1)} \tilde{\boldsymbol{c}}.$$
(30)

Once we have arrived at the representation in terms of orbitals, we can restrict the orbital basis to a non-complete set. This simply gives the operator in the variational (Fock) subspace spanned by the orbitals.

We note that the expression (29) not only works for local operators but also for differential operators like the momentum or kinetic energy: we have taken care not to exchange the order of  $M^{(1)}$  and one of its field operators. We can write truly non-local operators in a similar way. As an example, the one-body density operator is given by

$$\hat{\Gamma}^{(1)}(x;x') = \hat{\Psi}^{\dagger}(x)\hat{\Psi}(x')$$
(31)

so that one coordinate is not integrated over, rather setting it to x in the bra and x' in the ket. In an orthonormal basis it becomes

$$\hat{\Gamma}^{(1)}(x;x') = \sum_{n,m} \overline{\varphi_n(x)} \,\varphi_m(x') \,c_n^{\dagger} c_m \,. \tag{32}$$

#### 2.3.2 Two-body operators

For the two-body operators  $\sum_{i < j} M^{(2)}(x_i, x_j)$  we proceed in the familiar way, anti-commuting first the operators with the coordinates involved in  $M^{(2)}$  all the way to the left and right, respectively. This time we are left with a zero-body operator for N-2 electrons:

$$\hat{M}^{(2)} = \frac{1}{N!} \int dx_1 \cdots dx_N \hat{\Psi}^{\dagger}(x_N) \cdots \hat{\Psi}^{\dagger}(x_1) \sum_{i < j} M^{(2)}(x_i, x_j) \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N)$$

$$= \frac{1}{N!} \sum_{i < j} \int dx_i dx_j \hat{\Psi}^{\dagger}(x_j) \hat{\Psi}^{\dagger}(x_i) M^{(2)}(x_i, x_j) (N-2)! \hat{\Psi}(x_i) \hat{\Psi}(x_j)$$

$$= \frac{1}{N(N-1)} \sum_{i < j} \int dx_i dx_j \hat{\Psi}^{\dagger}(x_j) \hat{\Psi}^{\dagger}(x_i) M^{(2)}(x_i, x_j) \hat{\Psi}(x_i) \hat{\Psi}(x_j)$$

$$= \frac{1}{2} \int dx \, dx' \, \hat{\Psi}^{\dagger}(x') \, \hat{\Psi}^{\dagger}(x) \, M^{(2)}(x, x') \, \hat{\Psi}(x) \, \hat{\Psi}(x')$$

Expanding in an orthonormal basis, we get

$$\hat{M}^{(2)} = \frac{1}{2} \sum_{n,n',m,m'} \int dx dx' \overline{\varphi_{n'}(x')\varphi_n(x)} M^{(2)}(x,x') \varphi_m(x)\varphi_{m'}(x') \quad c_{n'}^{\dagger} c_n^{\dagger} c_m c_{m'}$$
$$= \frac{1}{2} \sum_{n,n',m,m'} \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle \qquad c_{n'}^{\dagger} c_n^{\dagger} c_m c_{m'} \tag{33}$$

where the exchange of the indices in the second line is a consequence of the way the Dirac state for two electrons is usually written: first index for the first coordinate, second index for the second, while taking the adjoint of the operators changes their order.  $M_{nn',mm'} = \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle$  transforms like a fourth-order tensor: Transforming to a different basis (13) gives

$$\tilde{M}_{\nu\nu',\mu\mu'}^{(2)} = \sum_{n,n',m,m'} U_{\nu n}^{\dagger} U_{\nu'n'}^{\dagger} M_{nn',mm'} U_{m\mu} U_{m'\mu'} \,.$$
(34)

Form the symmetry of the two-body operator  $M^{(2)}(x, x') = M^{(2)}(x', x)$  follows  $M_{nn', mm'} = M_{n'n, m'm}$ . Moreover,  $M_{nn, mm'}$  will not contribute to  $\hat{M}^{(2)}$  since  $c_n^{\dagger} c_n^{\dagger} = \{c_n^{\dagger}, c_n^{\dagger}\}/2 = 0$ , and likewise for  $M_{nn', mm}$ .

Note that the representation (33) is not quite as efficient as it could be: The terms with n and n' and/or m and m' exchanged connect the same basis states. Collecting these terms by introducing an ordering of the operators and using the symmetry of the matrix elements we obtain

$$\hat{M}^{(2)} = \sum_{n' > n, \ m' > m} c_{n'}^{\dagger} c_{n}^{\dagger} \underbrace{\left(M_{nn', \ mm'}^{(2)} - M_{n'n, \ mm'}^{(2)}\right)}_{=:\breve{M}_{nn', \ mm'}^{(2)}} c_{m} c_{m'} \,. \tag{35}$$

Since the states  $\{c_{n'}^{\dagger}c_{n}^{\dagger}|0\rangle | n' > n\}$  form a basis of the two-electron Hilbert space, considering nn' as the index of a basis state, the  $\breve{M}_{nn',mm'}^{(2)}$  form a two-electron matrix  $\breve{M}^{(2)}$ .

The procedure of rewriting operators in second quantization obviously generalizes to observables acting on more than two electrons in the natural way. We note that, while we started from a form of the operators (24) that was explicitly formulated in an N-electron Hilbert space, the results (25), (29), and (33) are of the same form no matter what value N takes. Thus these operators are valid not just on some N-electron Hilbert space, but on the entire Fock space. This is a particular strength of the second-quantized formalism.

### 2.4 Transforming the orbital basis

We noted in (13) that the creators transform in the same way as the orbitals they represent

$$|\beta_i\rangle = U|\alpha_i\rangle = \sum_j |\beta_j\rangle \langle \alpha_j |\alpha_i\rangle = \sum_\mu |\alpha_\mu\rangle \underbrace{\langle \alpha_\mu | U | \alpha_i\rangle}_{=:U_{\mu i}} \quad \rightsquigarrow \quad c^{\dagger}_{\beta_i} = \sum_\mu c^{\dagger}_{\alpha_\mu} U_{\mu i} \,, \tag{36}$$

so the "operators" really transform like states. Writing the transformation matrix as  $U = e^M$ , where M is anti-Hermitian,  $M^{\dagger} = -M$  when U is unitary, but can be any matrix when U is merely invertible, we can write the basis transformation in a form appropriate for operators:

$$c_{\beta_i}^{\dagger} = e^{\mathbf{c}^{\dagger} \mathbf{M} \mathbf{c}} c_{\alpha_{\mu}}^{\dagger} e^{-\mathbf{c}^{\dagger} \mathbf{M} \mathbf{c}} \,. \tag{37}$$

To see this, we use the Baker-Campbell-Hausdorff formula in the form

$$e^{\lambda A}B e^{-\lambda A} = B + \lambda [A, B] + \frac{\lambda^2}{2!} [A, [A, B]] + \frac{\lambda^3}{3!} [A, [A, [A, B]]] + \cdots,$$
 (38)

where the expansion coefficients follow by taking the derivatives of the left hand side at  $\lambda = 0$ , together with the commutator

$$[c^{\dagger}_{\alpha_{\mu}}c_{\alpha_{\nu}}, c^{\dagger}_{\alpha_{\kappa}}] = c^{\dagger}_{\alpha_{\mu}} \,\delta_{\nu,\kappa} \tag{39}$$

from which we obtain for the repeated commutators

$$\left[\sum_{\mu,\nu} M_{\mu\nu} c^{\dagger}_{\alpha_{\mu}} c_{\alpha_{\nu}}, \sum_{\kappa} c^{\dagger}_{\alpha_{\kappa}} \left(M^{n}\right)_{\kappa i}\right] = \sum_{\mu\nu\kappa} c^{\dagger}_{\alpha_{\mu}} M_{\mu\nu} \,\delta_{\nu,\kappa} \left(M^{n}\right)_{\kappa i} = \sum_{\mu} c^{\dagger}_{\alpha_{\mu}} \left(M^{n+1}\right)_{\mu i}.$$
 (40)

To keep the derivation simple, we have chosen to transform an operator from the orthonormal basis that we also used to write the exponential operator. Being linear, the transform works, of course, the same for an arbitrary creation operator.

Using this form of the basis transformation and noticing that  $e^{-c^{\dagger}Mc}|0\rangle = |0\rangle$ , we immediately see that acting with the exponential of a one-body operator on a product state results in another product state

$$e^{\mathbf{c}^{\dagger}\mathbf{M}\mathbf{c}}\prod c_{\alpha_{n}}^{\dagger}|0\rangle = \prod e^{\mathbf{c}^{\dagger}\mathbf{M}\mathbf{c}}c_{\alpha_{n}}^{\dagger}e^{-\mathbf{c}^{\dagger}\mathbf{M}\mathbf{c}}|0\rangle = \prod c_{\beta_{n}}^{\dagger}|0\rangle.$$
(41)

This is, e.g., used when working in the interaction picture. Anticommutators with transformed operators, (36), are simply  $\{c_{\alpha_j}, e^{c^{\dagger}Mc} c_{\alpha_i}^{\dagger} e^{-c^{\dagger}Mc}\} = \langle \alpha_j | e^M | \alpha_i \rangle$ .

Annihilation operators, being the adjoint of the creators, transform in just the expected way

$$c_{\beta_i} = e^{-\boldsymbol{c}^{\dagger} \boldsymbol{M}^{\dagger} \boldsymbol{c}} c_{\alpha_{\mu}} e^{\boldsymbol{c}^{\dagger} \boldsymbol{M}^{\dagger} \boldsymbol{c}}, \qquad (42)$$

which means that for unitary transformations, where M is anti-Hermitian, creators and annihilators transform in the same way. Note that in the imaginary-time formalism the annihilators are, via analytic continuation, chosen to transform in the same way as the creators, making them different from the adjoint of the creators.

# **3** Variational methods

The variational principle and the Schrödinger equation are equivalent. Consider the energy expectation value as a wave-function functional

$$E[\Psi] = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} \,. \tag{43}$$

Its variation is

$$E[\Psi + \delta\Psi] = E[\Psi] + \frac{\langle \delta\Psi | H | \Psi \rangle + \langle \Psi | H | \delta\Psi \rangle}{\langle \Psi | \Psi \rangle} - \langle \Psi | H | \Psi \rangle \frac{\langle \delta\Psi | \Psi \rangle + \langle \Psi | \delta\Psi \rangle}{\langle \Psi | \Psi \rangle^2} + \mathcal{O}^2.$$
(44)

The first-order term vanishes for  $H|\Psi\rangle = E[\Psi] |\Psi\rangle$ , which is the Schrödinger equation. Since the eigenfunctions

$$H|\Psi_n\rangle = E_n|\Psi_n\rangle, \qquad (45)$$

can be chosen to form an orthonormal basis, we can expand any wavefunction as

$$|\Psi\rangle = \sum_{n} |\Psi_{n}\rangle \langle \Psi_{n}|\Psi\rangle \tag{46}$$

and determine, as long as  $\langle \Psi | \Psi \rangle \neq 0$ , its energy expectation value

$$\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{m,n} \langle \Psi | \Psi_m \rangle \langle \Psi_m | H | \Psi_n \rangle \langle \Psi_n | \Psi \rangle}{\sum_{m,n} \langle \Psi | \Psi_m \rangle \langle \Psi_m | \Psi_n \rangle \langle \Psi_n | \Psi \rangle} = \frac{\sum_n E_n \left| \langle \Psi_n | \Psi \rangle \right|^2}{\sum_n \left| \langle \Psi_n | \Psi \rangle \right|^2} \,. \tag{47}$$

Since by definition no eigenenergy can be lower than the ground state energy  $E_0$ , we immediately see that the energy expectation value can never drop below the ground state energy

$$\frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle} = \frac{\sum_{n} E_{n} \left| \langle \Psi_{n} | \Psi \rangle \right|^{2}}{\sum_{n} \left| \langle \Psi_{n} | \Psi \rangle \right|^{2}} \ge \frac{\sum_{n} E_{0} \left| \langle \Psi_{n} | \Psi \rangle \right|^{2}}{\sum_{n} \left| \langle \Psi_{n} | \Psi \rangle \right|^{2}} = E_{0}.$$
(48)

We can use the same argument to generalize this variational principle: Assume we have arranged the eigenenergies in ascending order,  $E_0 \leq E_1 \leq \cdots$ , then the energy expectation value for a wavefunction that is orthogonal to the *n* lowest eigenstates, can not drop below  $E_n$ 

$$\frac{\langle \Psi_{\perp_n} | H | \Psi_{\perp_n} \rangle}{\langle \Psi_{\perp_n} | \Psi_{\perp_n} \rangle} \ge E_n \quad \text{if } \langle \Psi_i | \Psi_{\perp_n} \rangle = 0 \text{ for } i = 0, \dots, n-1.$$
(49)

This generalized variational principle is, of course, only of practical use if we know something about the eigenstates, e.g., when we can use symmetries to ensure orthogonality.

For an ab-initio Hamiltonian of N electrons in the field of nuclei of charge  $Z_{\alpha}$  at position  $R_{\alpha}$ ,

$$H = -\frac{1}{2} \sum_{i} \Delta_{i} - \sum_{i,\alpha} \frac{Z_{\alpha}}{|\boldsymbol{r}_{i} - \boldsymbol{R}_{\alpha}|} + \sum_{i < j} \frac{1}{|\boldsymbol{r}_{i} - \boldsymbol{r}_{j}|} + \sum_{\alpha < \beta} \frac{Z_{\alpha} Z_{\beta}}{|\boldsymbol{R}_{\alpha} - \boldsymbol{R}_{\beta}|}, \quad (50)$$

the Schrödinger equation is a partial differential equation. In second quantization it becomes a linear-algebra problem: We introduce an orbital basis set  $\{\varphi_k \mid k\}$ , which for simplicity we

assume here to be orthonormal, from which we construct an orthonormal basis of N-electron product states,  $\{\Phi_{k_1,\ldots,k_N} | k_1 < \cdots < k_N\}$ . To simplify the notation we sort the basis states, e.g., lexicographically in the orbital indices  $\mathbf{k} = (k_1, \ldots, k_N)$  and define the row vector of basis states  $|\boldsymbol{\Phi}\rangle := (|\Phi_1\rangle, |\Phi_2\rangle, \ldots)$ . The expansion of a state  $|\Psi\rangle$  in this basis can then be written as

$$|\Psi\rangle = \sum_{k_1 < \dots < k_N} a_{k_1,\dots,k_N} |\Phi_{k_1,\dots,k_N}\rangle = \sum_i a_i |\Phi_i\rangle = |\Phi\rangle \,\boldsymbol{a} \,, \tag{51}$$

where a is the vector of expansion coefficients. Likewise we can write the Schrödinger equation as a matrix eigenvalue problem

$$\boldsymbol{H}\boldsymbol{a} = \langle \boldsymbol{\Phi} | \hat{H} | \boldsymbol{\Phi} \rangle \boldsymbol{a} = \begin{pmatrix} \langle \Phi_1 | \hat{H} | \Phi_1 \rangle & \langle \Phi_1 | \hat{H} | \Phi_2 \rangle \cdots \\ \langle \Phi_2 | \hat{H} | \Phi_1 \rangle & \langle \Phi_2 | \hat{H} | \Phi_2 \rangle \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E \begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E \boldsymbol{a} .$$
(52)

From the eigenvectors of the matrix H we easily recover the eigenstates of the Hamiltonian

$$\boldsymbol{H}\boldsymbol{a}_{n} = E_{n}\boldsymbol{a}_{n} \quad \rightsquigarrow \quad \hat{H}|\Psi_{n}\rangle = E_{n}|\Psi_{n}\rangle \quad \text{with } |\Psi_{n}\rangle = |\boldsymbol{\Phi}\rangle \,\boldsymbol{a}_{n} \,. \tag{53}$$

Unfortunately, for an ab-initio Hamiltonian like (50) we need an infinite orbital basis set, so that the Hamiltonian matrix  $\boldsymbol{H}$  is infinite dimensional. A pragmatic approach to allow for computer simulations is to simply restrict the calculation to a finite basis  $|\tilde{\boldsymbol{\Phi}}\rangle := (|\tilde{\Phi}_1\rangle, \dots, |\tilde{\Phi}_{\tilde{L}}\rangle)$ , i.e., work with a finite matrix  $\tilde{\boldsymbol{H}} := \langle \tilde{\boldsymbol{\Phi}} | \hat{H} | \tilde{\boldsymbol{\Phi}} \rangle$  of dimension  $\tilde{L}$ . The crucial question is then how the eigenstates

$$\tilde{\boldsymbol{H}}\tilde{\boldsymbol{a}}_{n} = \tilde{E}_{n}\tilde{\boldsymbol{a}}_{n} \quad \rightsquigarrow \quad |\tilde{\boldsymbol{\Psi}}_{n}\rangle := |\tilde{\boldsymbol{\Phi}}\rangle\,\tilde{\boldsymbol{a}}_{n} \tag{54}$$

are related to those of H. The answer is surprisingly simple [7]: The eigenvalues of  $\hat{H}$ , ordered as  $\tilde{E}_0 \leq \tilde{E}_1 \leq \cdots \leq \tilde{E}_{\tilde{L}-1}$ , are variational with respect to those of H:

$$E_n \le \tilde{E}_n \quad \text{for } n \in \{0, \dots, \tilde{L}-1\}.$$
(55)

To show this, we construct a state in span  $(|\tilde{\Psi}_0\rangle, \ldots, |\tilde{\Psi}_n\rangle)$ , which by construction has an energy expectation value  $\leq \tilde{E}_n$ , that is orthogonal to the exact eigenstates  $|\Psi_0\rangle, \ldots, |\Psi_{n-1}\rangle$ , so that by the generalized variational principle its expectation value is  $\geq E_n$ . Being the non-zero solution of n-1 linear equations with n variables, such a state certainly exists, hence  $E_n \leq \tilde{E}_n$ .

The convergence of the matrix eigenvalues with increasing basis size is surprisingly regular. Let us extend our original basis of  $\tilde{L}$  states by an additional  $L-\tilde{L}$  states. Then, repeating the above argument with the *L*-dimensional problem taking the role of  $\hat{H}$ , we obtain (55) with  $E_n$  being the eigenvalues of the *L*-dimensional Hamiltonian matrix H. Since H now is finite, we can use the same argument for -H, obtaining

$$-E_{L-i} \le -\tilde{E}_{\tilde{L}-i} \quad \text{for } i \in \{1, \dots, \tilde{L}\}.$$
(56)

Taking the two inequalities together we obtain

$$E_n \le \tilde{E}_n \le E_{n+(L-\tilde{L})} \quad \text{for } n \in \{0, \dots, \tilde{L}-1\}.$$
(57)

For the special case  $L = \tilde{L} + 1$  of adding a single basis state, this is the Hylleraas-Undheim/Mac-Donald nesting property for eigenvalues in successive approximations

$$E_1 \le E_1 \le E_2 \le E_2 \le \dots \le E_L \le E_{L+1}.$$
(58)

For a finite basis sets of K single-electron functions the dimension of (52) for an N-electron problem increase extremely rapidly: There are  $K \cdot (K-1) \cdot (K-2) \cdots (K-(N-1))$  ways of picking N indices out of K. Since we only use one specific ordering of these indices, we still have to divide by N! to obtain the number of such determinants:

$$\dim \mathcal{H}_K^{(N)} = \frac{K!}{N!(K-N)!} = \binom{K}{N}.$$
(59)

For N = 25 electrons and K = 100 orbitals the dimension already exceeds  $10^{23}$ .

To get reliable results, we simply have to systematically increase the basis until the change in the desired eigenvalues becomes smaller than the accuracy required by the physical problem. The art is, of course, to devise clever basis sets such that this is achieved already for bases of manageable dimensions.

#### **3.1** Non-interacting electrons

Even when considering a system of N non-interacting electrons we have to solve the large matrix eigenvalue problem (52). Writing the non-interacting Hamiltonian in the basis used for the CI expansion (51) we obtain

$$\hat{H} = \sum_{n,m} H_{nm} \, c_n^{\dagger} c_m \,,$$

which, in general, has non-vanishing matrix elements between Slater determinants that differ in at most one operator. But we can simplify things drastically by realizing that we can choose any single-electron basis for the CI expansion. If we choose the eigenstates of the single-electron matrix  $H_{nm}$  as basis, the second-quantized Hamiltonian becomes

$$\hat{H} = \sum_{n,m} \left( \varepsilon_n \, \delta_{n,m} \right) c_n^{\dagger} c_m = \sum_n \varepsilon_n \, c_n^{\dagger} c_n \, .$$

In this basis all off-diagonal matrix elements vanish and the CI Hamiltonian (52) is diagonal. Thus all  $\binom{K}{N}$  eigenstates are Slater determinants

$$|\Phi_{n}\rangle = c_{n_{N}}^{\dagger} \cdots c_{n_{1}}^{\dagger}|0\rangle$$
 with eigenenergy  $E_{n} = \sum_{i} \varepsilon_{n_{i}}$ . (60)

This shows that choosing an appropriate basis for a CI expansion is crucial. A good general strategy should thus be to solve the matrix problem (52) and at the same time look for the basis set (of given size) that minimizes the variational energy. This is the idea of the multi-configurational self-consistent field method (MCSCF) [8]. In the following we will restrict ourselves to the simplest case where the many-body basis consists of a single Slater determinant. This is the Hartree-Fock method.

## **3.2 Hartree-Fock theory**

The idea of the Hartree-Fock approach is to find an approximation to the ground-state of the N-electron problem by minimizing the total-energy wave-function functional (43) allowing only N-electron Slater determinants as variational functions. Since expectation values of Slater determinants are determined by their one-body density matrix, this means that we want to find the occupied subspace for which (43) is minimized.

To perform these variations we represent unitary transformations of the orbital basis as in Sec. 2.4 (related to the Thouless representation of Slater determinants [9])

$$\hat{U}(\lambda) = e^{i\lambda\hat{M}}$$
 with  $\hat{M} = \sum_{\alpha,\beta} M_{\alpha\beta} c^{\dagger}_{\alpha} c_{\beta}$  hermitian. (61)

Using the Hausdorff expansion (38), the variation of the energy expectation value becomes

$$E(\lambda) = \langle \Phi | e^{i\lambda\hat{M}} \hat{H} e^{-i\lambda\hat{M}} | \Phi \rangle$$
  
=  $\langle \Phi | \hat{H} | \Phi \rangle + i\lambda \langle \Phi | [\hat{H}, \hat{M}] | \Phi \rangle + \frac{(i\lambda)^2}{2} \langle \Phi | [[\hat{H}, \hat{M}], \hat{M}] | \Phi \rangle + \cdots$  (62)

The energy functional is stationary for  $\Phi^{\rm HF}$  when

$$\langle \Phi^{\rm HF} | [\hat{H}, \, \hat{M}] | \Phi^{\rm HF} \rangle = 0 \tag{63}$$

for every hermitian single-electron operator  $\hat{M}$ . This condition is most easily understood when we work with orthonormal orbitals  $\{|\varphi\rangle|n\}$  from which the Slater determinant can be constructed:  $|\Phi^{\text{HF}}\rangle = c_N^{\dagger} \cdots c_1^{\dagger}|0\rangle$ . Then (63) is equivalent to

$$\langle \Phi^{\rm HF} | [\hat{H}, \, c_n^{\dagger} c_m + c_m^{\dagger} c_n] | \Phi^{\rm HF} \rangle = 0 \quad \forall \; n, \; m$$

(in fact,  $n \ge m$  suffices). Since

$$c_n^{\dagger} c_m | \Phi^{\rm HF} \rangle = \begin{cases} \delta_{n,m} | \Phi^{\rm HF} \rangle & \text{if } n, m \in \{1, \dots, N\} \\ 0 & \text{if } m \notin \{1, \dots, N\} \end{cases},$$

i.e., (63) is automatically fulfilled if both n and m are either occupied or unoccupied (virtual). This is not unexpected since transformations among the occupied or virtual orbitals, respectively, do not change the Slater determinant. The condition thus reduces to

$$\langle \Phi^{\rm HF} | c_m^{\dagger} c_n \hat{H} | \Phi^{\rm HF} \rangle = 0 \quad \forall \ m \in \{1, \dots, N\}, \ n \notin \{1, \dots, N\}.$$
(64)

In other words, for the Hamiltonian there are no matrix elements between the stationary Slater determinant and determinants that differ from it in one orbital. This condition that for the Hartree-Fock determinant the Hamiltonian does not produce single excitations is called the Brillouin theorem.

Let us consider a Hamiltonian with one- and two-body terms

$$\hat{H} = \sum_{n,m} c_n^{\dagger} T_{nm} c_m + \sum_{n > n', m > m'} c_n^{\dagger} c_{n'}^{\dagger} \left( U_{nn',mm'} - U_{nn',m'm} \right) c_{m'} c_m$$

Then for each  $n > N \ge m$  the singly-excited term

$$\left(T_{nm} + \sum_{m' \le N} \left(U_{nm',mm'} - U_{nm',m'm}\right)\right) c_n^{\dagger} c_m |\Phi^{\rm HF}\rangle = 0$$

must vanish. This is the same condition as for a non-interacting Hamiltonian with matrix elements

$$F_{nm} = T_{nm} + \sum_{m' \le N} (U_{nm',mm'} - U_{nm',m'm}).$$
(65)

F is called the Fock matrix. It depends, via the summation over occupied states, i.e., the density matrix, on the Slater determinant it is acting on. So we cannot simply diagonalize the single-electron matrix F since this will, in general, give a different determinant. Instead we need to find a Slater determinant for which F is diagonal (in fact, it is sufficient if it is block-diagonal in the occupied and virtual spaces). This is typically done by constructing a new Slater determinant from the N lowest eigenstates of F and iterating. Alternatively, we can use, e.g., steepest descent methods to minimize the expectation value directly or optimizing the one-body density matrix [10, 11]. At self-consistency the Fock matrix is diagonal with eigenvalues

$$\varepsilon_m^{\rm HF} = \left( T_{mm} + \sum_{m' \le N} \underbrace{\left( U_{mm',mm'} - U_{mm',m'm} \right)}_{=:\Delta_{mm'}} \right) = \left( T_{mm} + \sum_{m' \le N} \Delta_{mm'} \right) \tag{66}$$

and the Hartree-Fock energy is given by

$$\langle \Phi^{\rm HF} | \hat{H} | \Phi^{\rm HF} \rangle = \sum_{m \le N} \left( T_{mm} + \sum_{m' < m} \Delta_{mm'} \right) = \sum_{m \le N} \left( T_{mm} + \frac{1}{2} \sum_{m' \le N} \Delta_{mm'} \right).$$

Removing an electron from the occupied orbital  $\varphi_a$  changes the energy expectation value by

$$\langle \Phi_{a \, \text{rem}}^{\text{HF}} | \hat{H} | \Phi_{a \, \text{rem}}^{\text{HF}} \rangle - \langle \Phi^{\text{HF}} | \hat{H} | \Phi^{\text{HF}} \rangle = -\left( T_{aa} + \frac{1}{2} \sum_{m' \le N} \Delta_{am'} \right) - \frac{1}{2} \sum_{m \ne a \le N} \Delta_{ma} = -\varepsilon_a^{\text{HF}}.$$
(67)

When we assume that removing an electron does not change the orbitals, which should be a good approximation in the limit of many electrons  $N \gg 1$ , this gives the ionization energy (Koopmans' theorem). Likewise, the energy expectation value of an excited Slater determinant  $\Phi_{a \rightarrow b}^{\text{HF}}$  with an electron moved from orbital  $a \leq N$  to orbital b > N is

$$\varepsilon_{a\to b}^{\rm HF} = \langle \Phi_{a\to b}^{\rm HF} | \hat{H} | \Phi_{a\to b}^{\rm HF} \rangle - \langle \Phi^{\rm HF} | \hat{H} | \Phi^{\rm HF} \rangle = \varepsilon_b^{\rm HF} - \varepsilon_a^{\rm HF} - \Delta_{ab}$$
(68)

It can be interpreted as the energy of a state with an electron-hole excitation, again neglecting relaxation effects. For the Coulomb interaction

$$\begin{aligned} \Delta_{ab} &= \frac{1}{2} (\Delta_{ab} + \Delta_{ba}) = \frac{1}{2} \left( \left\langle \varphi_a \varphi_b \left| \frac{1}{r - r'} \right| \varphi_a \varphi_b - \varphi_b \varphi_a \right\rangle + \left\langle \varphi_b \varphi_a \left| \frac{1}{r - r'} \right| \varphi_b \varphi_a - \varphi_a \varphi_b \right\rangle \right) \\ &= \frac{1}{2} \left\langle \varphi_a \varphi_b - \varphi_b \varphi_a \left| \frac{1}{r - r'} \right| \varphi_a \varphi_b - \varphi_b \varphi_a \right\rangle > 0 \end{aligned}$$

so that the third term in (68) describes the attraction between the excited electron and the hole.

#### 3.2.1 Homogeneous electron gas

Since the homogeneous electron gas is translation invariant it is natural to write the Hamiltonian (for states with homogeneous charge density) in the basis of plane waves  $\langle r, \sigma | \mathbf{k}, \sigma \rangle = (2\pi)^{-3/2} e^{i\mathbf{k}\cdot\mathbf{r}}$ 

$$\hat{H} = \sum_{\sigma} \int d\boldsymbol{k} \, \frac{|\boldsymbol{k}^2|}{2} c^{\dagger}_{\boldsymbol{k},\sigma} c_{\boldsymbol{k},\sigma} + \frac{1}{2(2\pi)^3} \sum_{\sigma,\sigma'} \int d\boldsymbol{k} \int d\boldsymbol{k}' \int' d\boldsymbol{q} \, \frac{4\pi}{|\boldsymbol{q}|^2} c^{\dagger}_{\boldsymbol{k}-\boldsymbol{q},\sigma} c^{\dagger}_{\boldsymbol{k}'+\boldsymbol{q},\sigma'} c_{\boldsymbol{k}',\sigma'} c_{\boldsymbol{k},\sigma} \,, \quad (69)$$

where the prime on the q integral means that q = 0 is excluded since the homogeneous contribution to the Coulomb repulsion of the electrons is cancelled by its attraction with the homogeneous neutralizing background charge density. It seems reasonable to consider as an ansatz a Slater determinant  $|\Phi_{k_F}\rangle$  of all plane-wave states with momentum below some Fermi momentum,  $|\mathbf{k}| < k_F$ . The charge density for such a determinant follows, using the anticommutator of the field operator

$$\{\hat{\Psi}^{\dagger}_{\sigma}(\boldsymbol{r}), \, c_{\boldsymbol{k},\sigma}\} = \int d\boldsymbol{r}' \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}}{(2\pi)^{3/2}} \, \{\hat{\Psi}^{\dagger}_{\sigma}(\boldsymbol{r}), \, \hat{\Psi}_{\sigma}(\boldsymbol{r}')\} = \frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{r}}}{(2\pi)^{3/2}} \,,$$

from the diagonal of the density matrix

$$n_{\sigma}(\boldsymbol{r}) = \langle \Phi^{\rm HF} | \hat{\Psi}_{\sigma}^{\dagger}(\boldsymbol{r}) \hat{\Psi}_{\sigma}(\boldsymbol{r}) | \Phi^{\rm HF} \rangle = \int_{|\boldsymbol{k}| < k_F} d\boldsymbol{k} \left| \frac{e^{i\boldsymbol{k}\cdot\boldsymbol{r}}}{(2\pi)^{3/2}} \right|^2 = \frac{k_F^3}{6\pi^2}.$$
(70)

It is independent of position, so  $|\Phi_{k_F}\rangle$  looks like an appropriate ansatz for a homogeneous system. Moreover, it fulfills the stationarity condition (64): To create just a single excitation one of the creation operators in the Coulomb term of (69) must fill one of the annihilated states, i.e., q = 0 or q = k - k'. But this implies that the term is diagonal with q = 0 giving the direct and q = k - k' the exchange contribution. Since the q = 0 term is not present in the Hamiltonian, the eigenenergies of the Fock matrix are just the sum of the kinetic and the exchange terms

$$\varepsilon_{k,\sigma}^{\rm HF} = \frac{|\boldsymbol{k}|^2}{2} - \frac{1}{4\pi^2} \int_{|\boldsymbol{k}'| < k_F} d\boldsymbol{k}' \frac{1}{|\boldsymbol{k} - \boldsymbol{k}'|^2} = \frac{k^2}{2} - \frac{k_F}{\pi} \left( 1 + \frac{k_F^2 - k^2}{2k_F k} \ln \left| \frac{k_F + k}{k_F - k} \right| \right)$$
(71)

It depends only on  $k = |\mathbf{k}|$ . Interestingly the slope of  $\varepsilon_{\mathbf{k},\sigma}^{\text{HF}}$  becomes infinite for  $k \to k_F$ . Thus, the density of states  $D(\varepsilon_k)d\varepsilon = 4\pi k^2 dk$ , given by

$$D_{\sigma}^{\rm HF}(\varepsilon) = 4\pi k^2 \left(\frac{d\varepsilon_{k,\sigma}^{\rm HF}}{dk}\right)^{-1} = 4\pi k^2 \left(k - \frac{k_F}{\pi k} \left(1 - \frac{k_F^2 + k^2}{2k_F k} \ln\left|\frac{k_F + k}{k_F - k}\right|\right)\right)^{-1}$$
(72)

vanishes at the Fermi level (see Fig. 1). This is not quite what we expect from a respectable electron gas... It is clearly a defect of the Hartree-Fock approximation.

Instead of calculating the energy expectation value also directly in k-space, it is instructive to look at the exchange term in real space. To evaluate the electron-electron repulsion we need the diagonal of the 2-body density matrix, which is given in terms of the one-body density matrix

$$\langle \Phi_{k_F} | \hat{\Psi}^{\dagger}_{\sigma'}(\mathbf{r}') \hat{\Psi}^{\dagger}_{\sigma}(\mathbf{r}) \hat{\Psi}_{\sigma}(\mathbf{r}) \hat{\Psi}_{\sigma'}(\mathbf{r}') | \Phi_{k_F} 
angle = \det \begin{pmatrix} \Gamma^{(1)}_{\sigma\sigma}(\mathbf{r},\mathbf{r}) & \Gamma^{(1)}_{\sigma\sigma'}(\mathbf{r},\mathbf{r}') \\ \Gamma^{(1)}_{\sigma'\sigma}(\mathbf{r}',\mathbf{r}) & \Gamma^{(1)}_{\sigma'\sigma'}(\mathbf{r}',\mathbf{r}') \end{pmatrix},$$



**Fig. 1:** Hartree-Fock eigenvalues and density of states for the homogeneous solution  $|\Psi_{k_F}\rangle$  of the homogeneous electron gas compared to non-interacting values.

where the one-body density matrix vanishes unless  $\sigma' = \sigma$  where it is evaluated as in (70)

$$\Gamma_{\sigma\sigma}(\boldsymbol{r},\boldsymbol{r}') = \langle \Phi_{k_F} | \hat{\Psi}_{\sigma}^{\dagger}(\boldsymbol{r}) \hat{\Psi}_{\sigma}(\boldsymbol{r}') | \Phi_{k_F} \rangle \\
= \int_{|\boldsymbol{k}| < k_F} d\boldsymbol{k} \, \frac{e^{-i\boldsymbol{k}\cdot(\boldsymbol{r}-\boldsymbol{r}')}}{(2\pi)^3} = \frac{1}{4\pi^2} \int_0^{k_F} d\boldsymbol{k} \, k^2 \int_{-1}^1 d\cos\theta \, e^{i\boldsymbol{k}|\boldsymbol{r}-\boldsymbol{r}'|\cos\theta} \\
= \frac{k_F^3}{2\pi^2} \underbrace{\frac{\sin x - x\cos x}{x^3}}_{x \to 0 \ 1/3} = 3n_\sigma \, \frac{\sin x - x\cos x}{x^3}$$
(73)

with  $x = k_F |\mathbf{r} - \mathbf{r'}|$ . Dividing the 2-body density matrix by  $n_{\sigma}^2$  and subtracting the direct term (which is canceled by the contribution of the background charge) we obtain the exchange hole [11]

$$g_x(r,0) - 1 = -9\left(\frac{\sin k_F r - k_F r \cos k_F r}{(k_F r)^3}\right)^2.$$
 (74)

It is shown in Fig. 2. The exchange energy per spin is then the Coulomb interaction of the charge density with its exchange hole

$$E_x = \frac{1}{2} \int d\mathbf{r} \, n_\sigma \int d\mathbf{r}' n_\sigma \frac{g_x(r,r') - 1}{|\mathbf{r} - \mathbf{r}'|} = \frac{1}{2} \underbrace{\int d\mathbf{r} \, n_\sigma}_{=N} \int d\tilde{\mathbf{r}} \, n_\sigma \, \frac{g_x(\tilde{r},0) - 1}{\tilde{r}}$$

The exchange energy per electron of spin  $\sigma$  is thus

$$\varepsilon_x^{\sigma} = \frac{4\pi n_{\sigma}}{2} \int_0^\infty dr \, r^2 \, \frac{g(r,0) - 1}{r} = -\frac{9 \cdot 4\pi n_{\sigma}}{2k_F^2} \underbrace{\int_0^\infty dx \, \frac{(\sin x - x \cos x)^2}{x^5}}_{=1/4} = -\frac{3k_F}{4\pi} \,. \tag{75}$$



**Fig. 2:** Exchange hole for a paramagnetic homogeneous electron gas in units of the spin Wigner-Seitz radius  $k_F r_{\sigma} = (9\pi/2)^{1/3}$ . In addition, the dotted line shows the contribution of the exchange hole to the Coulomb repulsion energy of Eq. (75).

Together with the kinetic energy per electron of spin  $\sigma$ 

$$\varepsilon_{\rm kin}^{\sigma} = 4\pi \int_0^{k_F} dk \, k^2 \, \frac{k^2}{2} \, \bigg/ \, 4\pi \int_0^{k_F} dk \, k^2 = \frac{3k_F^2}{10} \tag{76}$$

we obtain the total energy per electron

$$\varepsilon_{\rm HF} = \frac{n_{\uparrow} \left(\varepsilon_{\rm kin}^{\uparrow} + \varepsilon_x^{\uparrow}\right) + n_{\downarrow} \left(\varepsilon_{\rm kin}^{\downarrow} + \varepsilon_x^{\downarrow}\right)}{n_{\uparrow} + n_{\downarrow}} = \frac{3(6\pi^2)^{2/3}}{10} \frac{n_{\uparrow}^{5/3} + n_{\downarrow}^{5/3}}{n} - \frac{3}{4} \left(\frac{6}{\pi}\right)^{1/3} \frac{n_{\uparrow}^{4/3} + n_{\downarrow}^{4/3}}{n}$$

While the kinetic energy is lowest when  $n_{\uparrow} = n_{\downarrow}$ , exchange favors spin polarization. For reasonable electron densities the kinetic energy dominates, only at extremely low densities exchange dominates and the solution would be ferromagnetic.

A ferromagnetic Slater determinant would, of course, have two different Fermi momenta,  $k_F^{\uparrow} \neq k_F^{\downarrow}$ . It also would break the symmetry of the Hamiltonian under spin rotations. This is an example of how we can lower the energy expectation value by allowing Slater determinants that break a symmetry of the system. When we do not restrict the symmetry of the Slater determinant, the approach is called unrestricted Hartree-Fock. For the electron gas this approach actually gives Hartree-Fock states that even break translational symmetry, see, e.g., [12].

#### 3.2.2 Hubbard model

As a simple example to illustrate the difference between restricted and unrestricted Hartree-Fock we consider the Hubbard model with two sites, i = 1, 2, between which the electrons can hop with matrix element -t and with an on-site Coulomb repulsion U

$$\hat{H} = -t \sum_{\sigma} \left( c_{2\sigma}^{\dagger} c_{1\sigma} + c_{1\sigma}^{\dagger} c_{2\sigma} \right) + U \sum_{i \in \{1,2\}} n_{i\uparrow} n_{i\downarrow} \,. \tag{77}$$

The number of electrons N and the total spin projection  $S_z$  are conserved, so the Fock space Hamiltonian is block-diagonal in the Hilbert spaces with fixed number of up- and down-spin electrons  $N_{\uparrow}$  and  $N_{\downarrow}$  with dimensions

N	0	1		2			3		4		
$N_{\uparrow}$	0	1	0	2	1	0	2	1	2		
$N_{\downarrow}$	0	0	1	0	1	2	1	2	2		
dim	1	2	2	1	4	1	2	2	1	16	

**Exact solutions:** The Hamiltonian for  $N = N_{\uparrow} = 1$  is easily constructed. By introducing the basis states  $c^{\dagger}_{1\uparrow}|0\rangle$  and  $c^{\dagger}_{2\uparrow}|0\rangle$ , we obtain the Hamiltonian matrix

$$\left\langle 0 \middle| \begin{pmatrix} c_{1\uparrow} \\ c_{2\uparrow} \end{pmatrix} \hat{H} \left( c_{1\uparrow}^{\dagger} & c_{2\uparrow}^{\dagger} \right) \middle| 0 \right\rangle = \left( \begin{array}{cc} 0 & -t \left\langle 0 \middle| c_{1\uparrow} & c_{1\uparrow}^{\dagger} c_{2\uparrow} & c_{2\uparrow}^{\dagger} \middle| 0 \right\rangle \\ -t \left\langle 0 \middle| c_{2\uparrow} & c_{2\uparrow}^{\dagger} c_{1\uparrow} & c_{1\uparrow}^{\dagger} \middle| 0 \right\rangle & 0 \end{array} \right) = \left( \begin{array}{c} 0 & -t \\ -t & 0 \end{array} \right).$$

This is easily diagonalized giving the familiar bonding and antibonding solution

$$|\varphi_{\pm}\rangle = \frac{1}{\sqrt{2}} \left( c_{1\uparrow}^{\dagger} \pm c_{2\uparrow}^{\dagger} \right) |0\rangle = c_{\pm\uparrow}^{\dagger} |0\rangle .$$
(78)

For  $N_{\uparrow} = 1 = N_{\downarrow}$ , we obtain a non-trivial interacting system

$$\left\langle 0 \middle| \begin{pmatrix} c_{1\uparrow}c_{2\downarrow} \\ c_{2\uparrow}c_{1\downarrow} \\ c_{1\uparrow}c_{1\downarrow} \\ c_{2\uparrow}c_{2\downarrow} \end{pmatrix} \hat{H} \left( c_{2\downarrow}^{\dagger}c_{1\uparrow}^{\dagger} \quad c_{1\downarrow}^{\dagger}c_{2\uparrow}^{\dagger} \quad c_{1\downarrow}^{\dagger}c_{1\uparrow}^{\dagger} \quad c_{2\downarrow}^{\dagger}c_{2\uparrow}^{\dagger} \right) \middle| 0 \right\rangle = \begin{pmatrix} 0 & 0 & -t & -t \\ 0 & 0 & -t & -t \\ -t & -t & U & 0 \\ -t & -t & 0 & U \end{pmatrix}.$$
(79)

To diagonalize the matrix, we transform the basis into linear combinations of covalent and ionic states

$$|\text{cov}_{\pm}\rangle = \frac{1}{\sqrt{2}} \left( c_{2\downarrow}^{\dagger} c_{1\uparrow}^{\dagger} \pm c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} \right) |0\rangle \tag{80}$$

$$|\mathrm{ion}_{\pm}\rangle = \frac{1}{\sqrt{2}} \left( c_{1\downarrow}^{\dagger} c_{1\uparrow}^{\dagger} \pm c_{2\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} \right) |0\rangle \tag{81}$$

It is then easy to verify that  $|cov_{-}\rangle$  is an eigenstate with eigenvalue  $\varepsilon_{cov_{-}} = 0$  and that  $|ion_{-}\rangle$  has eigenenergy  $\varepsilon_{ion_{-}} = U$ . The remaining two states mix

$$\begin{pmatrix} \langle \operatorname{cov}_{+} | \\ \langle \operatorname{ion}_{+} | \end{pmatrix} \hat{H} \left( | \operatorname{cov}_{+} \rangle \quad | \operatorname{ion}_{+} \rangle \right) = \frac{1}{2} \left( U - \begin{pmatrix} U & 4t \\ 4t & -U \end{pmatrix} \right).$$

$$(82)$$

Rewriting the matrix

$$\begin{pmatrix} U & 4t \\ 4t & -U \end{pmatrix} = \sqrt{U^2 + 16t^2} \begin{pmatrix} \cos \Theta & \sin \Theta \\ \sin \Theta & -\cos \Theta \end{pmatrix},$$
(83)



**Fig. 3:** Spectrum of the two-site Hubbard model as a function of U/t.

we find the ground state of the half-filled two-site Hubbard model

$$|\text{gs}\rangle = \cos\Theta/2 |\text{cov}_+\rangle + \sin\Theta/2 |\text{ion}_+\rangle$$
(84)

$$= \frac{1}{\sqrt{2}} \left( \cos \frac{\Theta}{2} c_{2\downarrow}^{\dagger} c_{1\uparrow}^{\dagger} + \cos \frac{\Theta}{2} c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} + \sin \frac{\Theta}{2} c_{1\downarrow}^{\dagger} c_{1\uparrow}^{\dagger} + \sin \frac{\Theta}{2} c_{2\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} \right) \left| 0 \right\rangle$$
(85)

with an energy of  $\varepsilon_{gs} = (U - \sqrt{U^2 + 16t^2})/2$ . Without correlations  $(U = 0 \rightarrow \Theta = \pi/2)$ , all basis states have the same prefactor, so we can factorize the ground state, writing it as a product  $c^{\dagger}_{+\downarrow}c^{\dagger}_{+\uparrow}|0\rangle$  of the operators defined in (78). For finite U this is no longer possible. In the strongly correlated limit  $U \gg t$  ( $\Theta \searrow 0$ ) the ground state becomes the maximally entangled state  $|cov_+\rangle$  and can not even approximately be expressed as a two-electron Slater determinant.

**Hartree-Fock:** We now want to see what Hartree-Fock can do in such a situation. Since the Hamiltonian is so simple, we can directly minimize the energy expectation value. The most general ansatz is a Slater determinant of an orbital  $\varphi(\theta_{\uparrow}) = \sin(\theta_{\uparrow}) \varphi_1 + \cos(\theta_{\uparrow}) \varphi_2$  for the spin-up, and  $\varphi(\theta_{\downarrow}) = \sin(\theta_{\downarrow}) \varphi_1 + \cos(\theta_{\downarrow}) \varphi_2$  for the spin-down electron:

$$\left|\Phi(\theta_{\uparrow},\theta_{\downarrow})\right\rangle = \left(\sin(\theta_{\downarrow})\,c_{1\downarrow}^{\dagger} + \cos(\theta_{\downarrow})\,c_{2\downarrow}^{\dagger}\right)\left(\sin(\theta_{\uparrow})\,c_{1\uparrow}^{\dagger} + \cos(\theta_{\uparrow})\,c_{2\uparrow}^{\dagger}\right)\left|0\right\rangle\,.\tag{86}$$

The energy expectation value as a function of the parameters  $\theta_{\sigma}$  is then

$$\langle \Phi(\theta_{\uparrow},\theta_{\downarrow}) | \hat{H} | \Phi(\theta_{\uparrow},\theta_{\downarrow}) \rangle = -2t \left( \sin \theta_{\uparrow} \sin \theta_{\downarrow} + \cos \theta_{\uparrow} \cos \theta_{\downarrow} \right) \left( \cos \theta_{\uparrow} \sin \theta_{\downarrow} + \sin \theta_{\uparrow} \cos \theta_{\downarrow} \right) + U \left( \sin^2 \theta_{\uparrow} \sin^2 \theta_{\downarrow} + \cos^2 \theta_{\uparrow} \cos^2 \theta_{\downarrow} \right) .$$

$$(87)$$

If the Slater determinant respects the symmetry of the molecule under the exchange of sites (mirror symmetry of the H<sub>2</sub> molecule), it follows that the Hartree-Fock orbitals for both spins are the bonding state  $\varphi_+$  ( $\theta = \pi/4$ ). This is the *restricted Hartree-Fock* solution. The corresponding energy is  $E(\pi/4, \pi/4) = -2t + U/2$ . The unrelaxed excited determinants are obtained



**Fig. 4:** Energy expectation value for a Slater determinant  $\Phi(\theta, \pi/2 - \theta)$  for U = 0, t, 2t, ..., 6t. When  $U \leq 2t$  the minimum is at  $\theta = \pi/4$ . This is the Hartree-Fock solution with the bonding orbitals  $\varphi_+$  occupied. For  $U \geq 2t$ ,  $\theta = \pi/4$  is still an extremal point (restricted Hartree-Fock solution), but an energy minimum is only attained when the symmetry is broken (unrestricted Hartree-Fock solution).

by replacing occupied orbitals  $\varphi_+$  with  $\varphi_-$ . Altogether we obtain the restricted Hartree-Fock spectrum

$$E(\pi/4, \pi/4) = -2t + U/2$$
  

$$E(\pi/4, -\pi/4) = U/2$$
  

$$E(-\pi/4, \pi/4) = U/2$$
  

$$E(-\pi/4, -\pi/4) = 2t + U/2$$
(88)

Comparing to the energy for a state with both electrons of the same spin (E = 0), we see that there is no spin-triplet, i.e., Hartree-Fock breaks the spin symmetry [13]. The states (88) are *spin-contaminated* [14]. Even worse, the Hartree-Fock ground state, and consequently all the states, are independent of U. The weight of the ionic states is always 1/2, leading to an increase of the energy with U/2.

To avoid this, we can allow the Hartree-Fock solution to break the symmetry of the molecule (*unrestricted Hartree-Fock*), putting, e.g., more of the up-spin electron in the orbital on site 1 and more of the down-spin electron in orbital 2. In an extended system this corresponds to an antiferromagnetic spin-density wave. For U < 2t this does not lead to a state of lower energy. For larger U, however, there is a symmetry-broken ground state

$$\Phi_{\rm UHF} = \Phi(\theta, \pi/2 - \theta) \text{ with } \theta(U) = \frac{\pi}{4} \pm \frac{1}{2} \arccos\left(\frac{2t}{U}\right).$$
(89)

Its energy is  $E_{\text{UHF}} = -2t^2/U$ . Still there is no triplet state (spin contamination) and, for  $U \to \infty$ , the overlap with the true singlet ground state goes to  $|\langle \Phi_{\text{UHF}} | \Psi_{-} \rangle|^2 = 1/2$ .

From Fig. 4 it might appear that there are just two degenerate unrestricted Hartree-Fock determinants. But, remembering that we can chose the spin quantization axis at will, we see that by rotating the spins by an angle  $\alpha$  about the axis  $\hat{n}$  (see App. A.2)

$$R_{\hat{n}}(\alpha) = e^{-i\hat{n}\cdot\vec{\sigma}\,\alpha/2} = \cos(\alpha/2) - i\sin(\alpha/2)\,\hat{n}\cdot\vec{\sigma}$$

we can produce a continuum of degenerate solutions  $\hat{R}_{\hat{n}}(\alpha) | \Phi_{\text{UHF}} \rangle$ . As an example we consider the state we obtain when we rotate the spin quantization axis from the  $\hat{z}$  into the  $\hat{x}$  direction

$$R_{\hat{y}}(-\pi/2) = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1\\ -1 & 1 \end{pmatrix}$$

which transforms the creation operators according to (13) as

$$\left(c_{i\uparrow}^{\dagger}, c_{i\downarrow}^{\dagger}\right) R_{\hat{y}}(-\pi/2) = \left(\frac{1}{\sqrt{2}} \left(c_{i\uparrow}^{\dagger} - c_{i\downarrow}^{\dagger}\right), \frac{1}{\sqrt{2}} \left(c_{i\uparrow}^{\dagger} + c_{i\downarrow}^{\dagger}\right)\right).$$

The determinant (86) thus transforms to

$$\hat{R}_{\hat{y}}(-\pi/2)|\Phi(\theta_{\uparrow},\theta_{\downarrow})\rangle = \frac{1}{2} \Big( s_{\downarrow} \big( c_{1\uparrow}^{\dagger} + c_{1\downarrow}^{\dagger} \big) + c_{\downarrow} \big( c_{2\uparrow}^{\dagger} + c_{2\downarrow}^{\dagger} \big) \Big) \Big( s_{\uparrow} \big( c_{1\uparrow}^{\dagger} - c_{1\downarrow}^{\dagger} \big) + c_{\uparrow} \big( c_{2\uparrow}^{\dagger} - c_{2\downarrow}^{\dagger} \big) \Big) |0\rangle$$
(90)

where we introduced the abbreviations  $s_{\sigma} = \sin \theta_{\sigma}$  and  $c_{\sigma} = \cos \theta_{\sigma}$ . Since the Hamiltonian (77) is invariant under spin rotations,  $\hat{R}_{\hat{y}}(-\pi/2) \hat{H} \hat{R}_{\hat{y}}^{\dagger}(-\pi/2) = \hat{H}$ , the energy expectation value of the rotated state is still given by (87).

Attractive Hubbard model For negative U allowing the spin orbitals to differ,  $\Phi(\theta, \pi/2-\theta)$ , does not lower the energy expectation value. The minimum is always obtained for the restricted Hartree-Fock determinant  $\Phi(\pi/4, \pi/4)$ . In fact, for the attractive Hubbard model rather than breaking spin symmetry, we should try to break the charge symmetry: For U < -2t the ansatz  $\Phi(\theta, \theta)$  minimizes the energy for the two states  $\theta(U) = \pi/4 \pm \arccos(-2t/U)$  with energy  $E(U) = 2t^2/U + U$ . Thus, the unrestricted Hartree-Fock ground state breaks the charge symmetry, i.e., is a charge-density wave state. On the other hand, looking back to (90) we see that  $\Phi(\theta, \theta)$  is invariant under the spin rotation. This is actually true for any  $\hat{R}_{\hat{n}}(\alpha)$  so that the unrestricted Hartree-Fock ground state of the attractive Hubbard model does not break spin symmetry.

It seems strange that for the attractive model we only find two unrestricted Hartree-Fock states, while for the repulsive model we have a continuum of states. To find the 'missing' states we consider a new kind of transformation that mixes creation and annihilation operators: When we exchange the role of the creation and annihilation operators for the up spins only, i.e.,

$$\tilde{c}_{i\uparrow}^{\dagger} = (-1)^{i} c_{i\uparrow} \quad \text{and} \quad \tilde{c}_{i\downarrow}^{\dagger} \to c_{i\downarrow}^{\dagger},$$
(91)

the Hamiltonian (77) transforms into a two-site Hubbard model with the sign of U reversed

$$\hat{H} = -t \sum_{\sigma} \left( \tilde{c}_{2\sigma}^{\dagger} \tilde{c}_{1\sigma} + \tilde{c}_{1\sigma}^{\dagger} \tilde{c}_{2\sigma} \right) - U \sum_{i \in \{1,2\}} \tilde{n}_{i\uparrow} \tilde{n}_{i\downarrow} + U(\tilde{n}_{1\downarrow} + \tilde{n}_{2\downarrow}) \,. \tag{92}$$

Let us see what happens to the Slater determinant (86) when we apply the same transformation. In doing this, we have to remember that the vacuum state must vanish when acted on with an annihilator. For  $|0\rangle$  this is no longer true for the transformed operators, but we can easily write down a state

$$\tilde{0}\rangle = c_{2\uparrow}^{\dagger}c_{1\uparrow}^{\dagger}|0\rangle \tag{93}$$

that behaves as a suitable vacuum state:  $\tilde{c}_{i\sigma}|\tilde{0}\rangle = 0$  and  $\langle \tilde{0}|\tilde{0}\rangle$ . We can then rewrite the transformed Slater determinant (86) as

$$\begin{split} |\tilde{\Phi}(\theta_{\uparrow},\theta_{\downarrow})\rangle &= \left(\sin(\theta_{\downarrow})\,\tilde{c}_{1\downarrow}^{\dagger} + \cos(\theta_{\downarrow})\,\tilde{c}_{2\downarrow}^{\dagger}\right) \left(\sin(\theta_{\uparrow})\,\tilde{c}_{1\uparrow}^{\dagger} + \cos(\theta_{\uparrow})\,\tilde{c}_{2\uparrow}^{\dagger}\right) |\tilde{0}\rangle \\ &= \left(\sin(\theta_{\downarrow})\,c_{1\downarrow}^{\dagger} + \cos(\theta_{\downarrow})\,c_{2\downarrow}^{\dagger}\right) \left(-\sin(\theta_{\uparrow})\,c_{1\uparrow} + \cos(\theta_{\uparrow})\,c_{2\uparrow}\right) c_{2\uparrow}^{\dagger}c_{1\uparrow}^{\dagger}|0\rangle \\ &= \left(\sin(\theta_{\downarrow})\,c_{1\downarrow}^{\dagger} + \cos(\theta_{\downarrow})\,c_{2\downarrow}^{\dagger}\right) \left(+\sin(\theta_{\uparrow})\,c_{2\uparrow}^{\dagger} + \cos(\theta_{\uparrow})\,c_{1\uparrow}^{\dagger}\right) |0\rangle \;. \end{split}$$

Thus, the transformation takes the unrestricted state  $|\Phi(\theta, \pi/2 - \theta)\rangle$  for the repulsive Hubbard model into the unrestricted state  $|\Phi(\theta, \theta)\rangle$  for the attractive Hubbard model. Transforming the rotated state (90) in the same way, we find something remarkable:

$$\begin{split} &\frac{1}{2} \left( s_{\downarrow} (\tilde{c}_{1\uparrow}^{\dagger} + \tilde{c}_{1\downarrow}^{\dagger}) + c_{\downarrow} (\tilde{c}_{2\uparrow}^{\dagger} + \tilde{c}_{2\downarrow}^{\dagger}) \right) \left( s_{\uparrow} (\tilde{c}_{1\uparrow}^{\dagger} - \tilde{c}_{1\downarrow}^{\dagger}) + c_{\uparrow} (\tilde{c}_{2\uparrow}^{\dagger} - \tilde{c}_{2\downarrow}^{\dagger}) \right) |\tilde{0}\rangle \\ &= \frac{1}{2} \left( s_{\downarrow} (-c_{1\uparrow} + c_{1\downarrow}^{\dagger}) + c_{\downarrow} (c_{2\uparrow} + c_{2\downarrow}^{\dagger}) \right) \left( s_{\uparrow} (-c_{1\uparrow} - c_{1\downarrow}^{\dagger}) + c_{\uparrow} (c_{2\uparrow} - c_{2\downarrow}^{\dagger}) \right) c_{2\uparrow}^{\dagger} c_{1\uparrow}^{\dagger} |0\rangle \\ &= \frac{1}{2} \left( (s_{\downarrow} c_{\uparrow} + c_{\downarrow} s_{\uparrow}) (c_{1\downarrow}^{\dagger} c_{1\uparrow}^{\dagger} + c_{2\downarrow}^{\dagger} c_{2\uparrow}^{\dagger}) |0\rangle + 2 \left( s_{\downarrow} s_{\uparrow} c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} + c_{\downarrow} c_{\uparrow} c_{2\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} \right) |0\rangle \\ &\quad + (s_{\downarrow} c_{\uparrow} - c_{\downarrow} s_{\uparrow}) \left( c_{2\downarrow}^{\dagger} c_{1\downarrow}^{\dagger} c_{2\uparrow}^{\dagger} c_{1\uparrow}^{\dagger} - 1 \right) |0\rangle \qquad \qquad \Big). \end{split}$$

The energy expectation value of this state is by construction the same as for the charge-density state. For  $\theta_{\downarrow} = \pi/2 - \theta_{\uparrow}$  the new state has a uniform density, but the wavefunction no longer has a well-defined particle number, i.e., it breaks particle number conservation. It is still a product state in the transformed operators and vacuum, but it is a state in Fock space. States of this type are the key to describing superconductivity.

#### **3.3 BCS theory**

Next we consider the BCS Hamiltonian

$$\hat{H}_{BCS} = \sum_{k\sigma} \varepsilon_{k} c^{\dagger}_{k\sigma} c_{k\sigma} - \sum_{kk'} G_{kk'} c^{\dagger}_{k\uparrow} c^{\dagger}_{-k\downarrow} c_{-k'\downarrow} c_{k'\uparrow}$$
(94)

with an attractive interaction between pairs of electrons of opposite spin and momentum (Cooper pairs). We now want to see if we can use the idea of product states in Fock space that we encountered for the attractive Hubbard model. To start, let us consider the determinant of plane-wave states that we used for the homogeneous electron gas  $|\Phi_{k_F}\rangle$ . Since all states with momentum below  $k_F$  are occupied, we have

$$c^{\dagger}_{k\sigma} | \Phi_{k_F} \rangle = 0$$
 for  $| \mathbf{k} | < k_F$  and  $c^{\dagger}_{k\sigma} | \Phi_{k_F} \rangle = 0$  otherwise.

Thus  $|\Phi_{k_F}\rangle$  behaves like a vacuum state for the transformed operators

$$\tilde{c}_{\boldsymbol{k}\sigma} = \Theta(k_F - |\boldsymbol{k}|) c_{\boldsymbol{k}\sigma}^{\dagger} + \Theta(|\boldsymbol{k}| - k_F) c_{\boldsymbol{k}\sigma} = \begin{cases} c_{\boldsymbol{k}\sigma}^{\dagger} & \text{for } |\boldsymbol{k}| < k_F \\ c_{\boldsymbol{k}\sigma} & \text{for } |\boldsymbol{k}| > k_F \end{cases}$$

Allowing the operators to mix, we can generalize this transformation to

$$\begin{split} b_{\mathbf{k}\uparrow} &= u_k c_{\mathbf{k}\uparrow} - v_k c_{-\mathbf{k}\downarrow}^{\dagger} \\ b_{\mathbf{k}\downarrow} &= u_k c_{\mathbf{k}\downarrow} + v_k c_{-\mathbf{k}\uparrow}^{\dagger} \end{split}$$

Notice how states with  $(\mathbf{k}, \sigma)$  and  $(-\mathbf{k}, -\sigma)$  are mixed. The corresponding creation operators are the adjoints,  $b_{k\sigma}^{\dagger} = \overline{u}_k c_{k\sigma}^{\dagger} - 2\sigma \overline{v}_k c_{-k,-\sigma}$ , where  $\sigma = \pm 1/2$ . These Bogoliubov-Valatin operators fulfill the canonical anticommutation relations

$$\{b_{\boldsymbol{k}\sigma}, \, b_{\boldsymbol{k}'\sigma'}\} = 0 = \{b_{\boldsymbol{k}\sigma}^{\dagger}, \, b_{\boldsymbol{k}'\sigma'}^{\dagger}\} \quad \text{and} \quad \{b_{\boldsymbol{k}\sigma}, \, b_{\boldsymbol{k}'\sigma'}^{\dagger}\} = \delta(\boldsymbol{k}-\boldsymbol{k}')\,\delta_{\sigma,\sigma'}$$

when (the non-trivial anticommutators are  $\{b_{k\uparrow}, b_{-k\downarrow}\}$  and  $\{b_{k\sigma}, b_{k\sigma}^{\dagger}\}$ )

$$|u_k|^2 + |v_k|^2 = 1. (95)$$

A vacuum state for the new operators can be constructed from the generalized product state  $\prod_{k\sigma} b_{k\sigma} |0\rangle$ . Expanding the operators

$$b_{-\mathbf{k}\uparrow}b_{\mathbf{k}\downarrow}b_{\mathbf{k}\uparrow}b_{-\mathbf{k}\downarrow}|0\rangle = v_k(u_k + v_k c^{\dagger}_{-\mathbf{k}\uparrow}c^{\dagger}_{\mathbf{k}\downarrow}) v_k(u_k + v_k c^{\dagger}_{\mathbf{k}\uparrow}c^{\dagger}_{-\mathbf{k}\downarrow}) |0\rangle$$

and calculating the norm

$$\langle 0|(\overline{u}_k+\overline{v}_k c_{-\boldsymbol{k}\downarrow}c_{\boldsymbol{k}\uparrow})(\overline{u}_k+\overline{v}_k c_{\boldsymbol{k}\downarrow}c_{-\boldsymbol{k}\uparrow})(u_k+v_k c_{-\boldsymbol{k}\uparrow}^{\dagger}c_{\boldsymbol{k}\downarrow}^{\dagger})(u_k+v_k c_{\boldsymbol{k}\uparrow}^{\dagger}c_{-\boldsymbol{k}\downarrow}^{\dagger})|0\rangle = \left(|u_k|^2+|v_k|^2\right)^2$$

we see from (95) that the BCS wavefunction

$$|BCS\rangle = \prod_{k} (u_k + v_k c^{\dagger}_{k\uparrow} c^{\dagger}_{-k\downarrow}) |0\rangle$$
(96)

is the (normalized) vacuum for the Bogoliubov-Valatin operators.

To calculate physical expectation values we express the electron operators as

$$\begin{aligned} c_{\mathbf{k}\uparrow} &= \overline{u}_k b_{\mathbf{k}\uparrow} + v_k b_{-\mathbf{k}\downarrow}^{\dagger} \\ c_{\mathbf{k}\downarrow} &= \overline{u}_k b_{\mathbf{k}\downarrow} - v_k b_{-\mathbf{k}\uparrow}^{\dagger} \end{aligned}$$

The expectation value for the occupation of a plane-wave state, e.g., is

$$\langle \mathrm{BCS} | \hat{n}_{\boldsymbol{k}\uparrow} | \mathrm{BCS} \rangle = \langle \mathrm{BCS} | (u_k b_{\boldsymbol{k}\uparrow}^{\dagger} + \overline{v}_k b_{-\boldsymbol{k}\downarrow}) (\overline{u}_k b_{\boldsymbol{k}\uparrow} + v_k b_{-\boldsymbol{k}\downarrow}^{\dagger}) | \mathrm{BCS} \rangle = |v_k|^2 = \langle \mathrm{BCS} | \hat{n}_{-\boldsymbol{k}\downarrow} | \mathrm{BCS} \rangle.$$

Unlike the electron gas Slater determinant  $|\Phi_{k_F}\rangle$ , where  $n_{k\sigma}$  is 1 below  $k_F$  and vanishes above, varying the parameter  $v_k$  in the BCS wavefunction allows us to get arbitrary momentum distributions  $\langle n_{k\sigma} \rangle$ . Since the BCS wavefunction has contributions in all particle sectors with an even number of electrons, there are also less-conventional expectation values, e.g.,

$$\langle \mathrm{BCS} | c_{\mathbf{k}\uparrow}^{\dagger} c_{-\mathbf{k}\downarrow}^{\dagger} | \mathrm{BCS} \rangle = \langle \mathrm{BCS} | (u_{k} b_{\mathbf{k}\uparrow}^{\dagger} + \overline{v}_{k} b_{-\mathbf{k}\downarrow}) (u_{k} b_{-\mathbf{k}\downarrow}^{\dagger} - \overline{v}_{k} b_{\mathbf{k},\uparrow}) | \mathrm{BCS} \rangle = \overline{v}_{k} u_{k} = \langle c_{-\mathbf{k}\downarrow} c_{\mathbf{k}\uparrow} \rangle.$$

When minimizing the energy expectation value, we have to introduce a chemical potential  $\mu$  that is chosen to give the desired number of particles  $N = \sum_{k\sigma} |v_k|^2$ . We get

$$\langle \text{BCS} | \hat{H} - \mu \hat{N} | \text{BCS} \rangle = \sum_{k\sigma} (\varepsilon_k - \mu) |v_k|^2 - \sum_{k,k'} G_{kk'} \, \overline{v}_k u_k v_{k'} \overline{u}_{k'} \,, \tag{97}$$

where the last term is real since  $G_{kk'} = G_{k'k}$ . Writing the Bogoliubov coefficients in polar coordinates,  $u_k = |u_k|e^{i\alpha_k}$  and  $v_k = |v_k|e^{i(\alpha_k + \varphi_k)}$ , we obtain

$$\langle \text{BCS}|\hat{H}-\mu\hat{N}|\text{BCS}\rangle = \sum (\varepsilon_{\boldsymbol{k}}-\mu)|v_{\boldsymbol{k}}|^2 - \sum G_{\boldsymbol{k}\boldsymbol{k}'}|v_{\boldsymbol{k}}||u_{\boldsymbol{k}}||v_{\boldsymbol{k}'}||u_{\boldsymbol{k}'}|\cos(\varphi_{\boldsymbol{k}'}-\varphi_{\boldsymbol{k}}), \quad (98)$$

which is minimized when the relative phase  $\varphi_k$  is independent of k, i.e., all Cooper pairs in (96) are *phase coherent* (while the  $\alpha_k$  merely contribute to the global phase)

$$|\mathrm{BCS}(\varphi)\rangle = \prod \left( |u_k| + |v_k| e^{i\varphi} c^{\dagger}_{\mathbf{k}\uparrow} c^{\dagger}_{-\mathbf{k}\downarrow} \right) |0\rangle.$$
<sup>(99)</sup>

Their overlap is  $\langle BCS(\varphi')|BCS(\varphi)\rangle = \prod_{k} (|u_{k}|^{2} + |v_{k}|^{2}e^{i(\varphi'-\varphi)})$  so that in the thermodynamic limit they are orthogonal for  $\varphi \neq \varphi'$ . For a fixed phase  $\varphi$  the particle number fluctuates  $\langle \hat{N}^{2} \rangle - \langle \hat{N} \rangle^{2} = 4 \sum_{k} |v_{k}|^{2} |u_{k}|^{2} < 2 \langle \hat{N} \rangle$ . Since the phase of a state with N/2 Cooper pairs is  $e^{i\varphi N/2}$ , we can project out states with fixed number of electrons by integrating over  $\varphi$ 

$$|\mathrm{BCS}_N\rangle = \frac{1}{\sqrt{2\pi}} \int_0^{2\pi} d\varphi \, e^{-i\varphi N/2} \, |\mathrm{BCS}(\varphi)\rangle.$$

It remains to minimize (97) with respect to the absolute values, written as  $|u_k| = \sin \Theta_k$  and  $|v_k| = \cos \Theta_k$  with  $\Theta_k \in [0, \pi/2)$ . Setting the derivative of

$$\langle \text{BCS}|\hat{H} - \mu \hat{N}|\text{BCS}\rangle = \sum_{\boldsymbol{k}\sigma} (\varepsilon_{\boldsymbol{k}} - \mu) \frac{1 + \cos 2\Theta_{\boldsymbol{k}}}{2} - \sum_{\boldsymbol{k}\boldsymbol{k}'} G_{\boldsymbol{k}\boldsymbol{k}'} \frac{\sin 2\Theta_{\boldsymbol{k}}}{2} \frac{\sin 2\Theta_{\boldsymbol{k}'}}{2}$$

with respect to  $\Theta_k$  to zero we find the system of equations

$$\tan 2\Theta_k = -\frac{\sum_{k'} G_{kk'} \sin(2\Theta_{k'})/2}{\varepsilon_k - \mu} = -\frac{\Delta_k}{\varepsilon_k - \mu} = \frac{\Delta_k}{\mu - \varepsilon_k}$$
(100)

where we have introduced the gap function  $\Delta_k$ . Using  $(\cos x)^2 = 1/(1+(\tan x)^2)$  we obtain the momentum distribution

$$\langle n_{\boldsymbol{k}\sigma} \rangle = |v_k|^2 = \frac{1}{2} \left( 1 + \cos 2\Theta_k \right) = \frac{1}{2} \left( 1 - \frac{\varepsilon_{\boldsymbol{k}} - \mu}{\sqrt{(\varepsilon_{\boldsymbol{k}} - \mu)^2 + \Delta_{\boldsymbol{k}}^2}} \right).$$
(101)

For  $\Delta_k \equiv 0$  this is just the step function of a Fermi gas, for finite  $\Delta$  the transition is more smooth. We still have to determine the parameters  $\mu$  and  $\Delta_k$ . The chemical potential is fixed by

$$N = \sum_{\boldsymbol{k}} 2|v_{\boldsymbol{k}}|^2 = \sum_{\boldsymbol{k}} \left( 1 - \frac{\varepsilon_{\boldsymbol{k}} - \mu}{\sqrt{(\varepsilon_{\boldsymbol{k}} - \mu)^2 + \Delta_{\boldsymbol{k}}^2}} \right)$$
(102)

while for  $\Delta_k$  we obtain from (100) and (101) the self-consistent gap equations

$$\Delta_{\mathbf{k}} = \sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'} \frac{\sin 2\Theta_{\mathbf{k}'}}{2} = \frac{1}{2} \sum_{\mathbf{k}'} G_{\mathbf{k}\mathbf{k}'} \tan 2\Theta_{\mathbf{k}'} \cos 2\Theta_{\mathbf{k}'} = \frac{1}{2} \sum_{\mathbf{k}'} \frac{G_{\mathbf{k}\mathbf{k}'}\Delta_{\mathbf{k}'}}{\sqrt{(\varepsilon_{\mathbf{k}'} - \mu)^2 + \Delta_{\mathbf{k}'}^2}}.$$
 (103)



Fig. 5: Quasi-electron energy and density of states for the BCS state with and without gap.

To see that  $\Delta$  is indeed a gap, consider the (unrelaxed) quasi-electron states ( $\varepsilon_k > \mu$ )

$$|\boldsymbol{k}\uparrow\rangle = \frac{1}{u_k}c^{\dagger}_{\boldsymbol{k}\uparrow}|\mathrm{BCS}\rangle = b^{\dagger}_{\boldsymbol{k}\uparrow}|\mathrm{BCS}\rangle.$$
 (104)

Adding an electron of momentum k destroys its Cooper pair, changing  $\langle n_{k\uparrow}+n_{k\downarrow}\rangle$  from  $2|v_k^2|$  to 1 and removing the interaction of the pair with all others:

$$\begin{aligned} \langle \boldsymbol{k} \uparrow | \hat{H} - \mu \hat{N} | \boldsymbol{k} \uparrow \rangle - \langle \text{BCS} | \hat{H} - \mu \hat{N} | \text{BCS} \rangle &= (\varepsilon_{\boldsymbol{k}} - \mu) \left( 1 - 2|v_{\boldsymbol{k}}|^2 \right) + 2\Delta_{\boldsymbol{k}} |v_{\boldsymbol{k}}| |u_{\boldsymbol{k}}| \\ &= (\varepsilon_{\boldsymbol{k}} - \mu) \left( -\cos 2\Theta_{\boldsymbol{k}} \right) + \Delta_{\boldsymbol{k}} \sin 2\Theta_{\boldsymbol{k}} = \sqrt{(\varepsilon_{\boldsymbol{k}} - \mu)^2 + \Delta^2}. \end{aligned}$$

For  $\Delta_k \equiv 0$  we recover Koopmans' Hartree-Fock result, while for  $\Delta_k > 0$  a gap opens around the Fermi level. Fig. 5 compares the quasi-electron dispersion and the corresponding density of states for the two cases, assuming, for simplicity, a k-independent gap  $\Delta_k \equiv \Delta$ .

# 4 Conclusion

We have seen that second quantization is a remarkably useful formalism. With just a few simple rules for the field operators and the corresponding vacuum, it converts dealing with manyelectron states to straightforward algebraic manipulations. Moreover it is naturally suited for performing calculations in variational spaces spanned by a finite basis of orbitals. But its advantages go beyond a mere simplification. By abstracting from the coordinate representation, it allows us to express many-body operators in a way that is independent of the number of electrons. Because of this it becomes possible to consider Fock-space wavefunctions which do not have a definite number of electrons. This allows us to consider unrestricted mean-field states that not only break spatial or spin symmetries but also particle conservation. This additional freedom allows us to extend the concept of a Slater determinant to product states in Fock space, an example of which is the BCS wavefunction.

# A Appendix

### A.1 Non-orthonormal basis

A general one-electron basis of functions  $|\chi_n\rangle$  will have an overlap matrix  $S_{nm} = \langle \chi_n | \chi_m \rangle$  that is positive definite (and hence invertible) and hermitian. The completeness relation is

$$\mathbb{1} = \sum_{k,l} |\chi_k\rangle (S^{-1})_{kl} \langle \chi_l | \,. \tag{105}$$

With it we can easily write the Schrödinger equation  $\hat{H}|v\rangle = \varepsilon |v\rangle$  in matrix form

$$\sum_{k} \underbrace{\langle \chi_{i} | H | \chi_{k} \rangle}_{=:H_{ik}} \underbrace{\sum_{l} (S^{-1})_{kl} \langle \chi_{l} | v \rangle \langle \chi_{i} | \hat{H} | v \rangle}_{=:v_{k}} = \varepsilon \langle \chi_{i} | v \rangle = \varepsilon \sum_{k} \underbrace{\langle \chi_{i} | \chi_{k} \rangle}_{=S_{ik}} \underbrace{\sum_{l} (S^{-1})_{kl} \langle \chi_{l} | v \rangle}_{=v_{k}} .$$
(106)

Collecting all components, this becomes the generalized eigenvalue problem  $Hv = \varepsilon Sv$ . From the solution v we can easily construct  $|v\rangle = \sum v_k |\chi_k\rangle$ . It is, however, often more convenient to have an orthonormal basis, so that we do not have to deal with the overlap matrices in the definition of the second quantized operators or the generalized eigenvalue problem.

To orthonormalize the basis  $\{|\chi_n\rangle\}$ , we need to find a basis transformation T such that

$$|\varphi_n\rangle := \sum_m |\chi_m\rangle T_{mn}$$
 with  $\langle \varphi_n | \varphi_m \rangle = \delta_{mn}$ . (107)

This implies that  $T^{\dagger}ST = 1$ , or equivalently  $S^{-1} = TT^{\dagger}$ . This condition does not uniquely determine T. In fact there are many orthonormalization techniques, e.g., Gram-Schmidt orthonormalization or Cholesky decomposition.

Usually we will have chosen the basis functions  $|\chi_n\rangle$  for a physical reason, e.g., atomic orbitals, so that we would like the orthonormal basis functions to be as close to the original basis as possible, i.e, we ask for the basis transformation T that minimizes

$$\sum_{n} \left\| |\varphi_{n}\rangle - |\chi_{n}\rangle \right\|^{2} = \sum_{n} \left\| \sum_{m} |\chi_{m}\rangle (T_{mn} - \delta_{mn}) \right\|^{2}$$
  
= Tr ( $\mathbf{T}^{\dagger} - \mathbb{1}$ )  $\mathbf{S}$  ( $\mathbf{T} - \mathbb{1}$ ) = Tr ( $\underbrace{\mathbf{T}^{\dagger} \mathbf{S} \mathbf{T}}_{=\mathbb{1}} - \mathbf{T}^{\dagger} \mathbf{S} - \mathbf{S} \mathbf{T} + \mathbf{S}$ ). (108)

Given an orthonormalization T, we can obtain any other orthonormalization  $\tilde{T}$  by performing a unitary transformation, i.e.,  $\tilde{T} = TU$ . Writing  $U = \exp(i\lambda M)$  with M a Hermitian matrix, we obtain the variational condition

$$0 \stackrel{!}{=} \operatorname{Tr} \left( +i\boldsymbol{M}\boldsymbol{T}^{\dagger}\boldsymbol{S} - i\boldsymbol{S}\boldsymbol{T}\boldsymbol{M} \right) = i\operatorname{Tr} \left( \boldsymbol{T}^{\dagger}\boldsymbol{S} - \boldsymbol{S}\boldsymbol{T} \right)\boldsymbol{M},$$
(109)

which is fulfilled for  $ST = T^{\dagger}S$ , i.e.,  $ST^2 = T^{\dagger}ST = 1$ . The second variation at  $T = S^{-1/2}$ 

$$\frac{1}{2}\operatorname{Tr}\left(\boldsymbol{M}^{2}\boldsymbol{S}^{1/2} + \boldsymbol{S}^{1/2}\boldsymbol{M}^{2}\right) > 0$$
(110)

is positive, since S and the square of the hermitian matrix M are both positive definite. Hence the Löwdin symmetric orthogonalization [15]

$$T_{\text{L\"owdin}} = S^{-1/2} \tag{111}$$

minimizes the modification of the basis vectors due to orthogonalization.

### A.2 Pauli matrices

The spin matrices were defined by Pauli [16] as

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

They are hermitian and unitary, so that  $\sigma_j^2 = \sigma_j \sigma_j^{\dagger} = 1$ . Moreover, det  $\sigma_j = -1$  and Tr  $\sigma_j = 0$ . By explicit multiplication we find  $\sigma_x \sigma_y = i\sigma_z$ , from which we get the more symmetric equation  $\sigma_x \sigma_y \sigma_z = i$ . These relations are cyclic, which is easily seen by repeatedly using  $\sigma_j^2 = 1$ 

$$\sigma_x \sigma_y = \mathrm{i} \sigma_z \stackrel{\cdot \sigma_z}{\rightsquigarrow} \sigma_x \sigma_y \sigma_z = \mathrm{i} \stackrel{\sigma_x \cdot}{\rightsquigarrow} \sigma_y \sigma_z = \mathrm{i} \sigma_x \stackrel{\cdot \sigma_x}{\rightsquigarrow} \sigma_y \sigma_z \sigma_x = \mathrm{i} \stackrel{\sigma_y \cdot}{\rightsquigarrow} \sigma_z \sigma_x = \mathrm{i} \sigma_y \stackrel{\cdot \sigma_y}{\rightsquigarrow} \sigma_z \sigma_x \sigma_y = \mathrm{i}.$$

Exchanging two adjacent indices changes the sign, e.g., multiplying  $\sigma_x \sigma_y \sigma_z = i$  from the left with  $\sigma_y \sigma_x$  gives  $\sigma_y \sigma_x = -i\sigma_z$ , which is again cyclic in the indices. We note that the multiplication table of the matrices  $-i\sigma_j$  is the same as the that of the basic quaternions. We can summarize the products of the Pauli matrices in the form

$$(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \sum a_j b_k \, \sigma_j \sigma_k = (\vec{a} \cdot \vec{b}) \mathbb{1} + i(\vec{a} \times \vec{b}) \cdot \vec{\sigma} \,.$$
 (112)

From the products follow the familiar commutation relations  $[\sigma_x, \sigma_y] = 2i\sigma_z$  (cyclic), while the anticommutators are  $\{\sigma_j, \sigma_k\} = 2\delta_{j,k} \mathbb{1}$ .

Together with the unit matrix, the Pauli matrices form a basis of the four-dimensional algebra of complex  $2 \times 2$  matrices and we can write

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = M = m_0 \mathbb{1} + \vec{m} \cdot \vec{\sigma} = \begin{pmatrix} m_0 + m_z & m_x - im_y \\ m_x + im_y & m_0 - m_z \end{pmatrix}$$
(113)

with  $2m_0 = m_{11} + m_{22}$ ,  $2m_z = m_{11} - m_{22}$ ,  $2m_x = m_{12} + m_{21}$ , and  $2m_x = i(m_{12} - m_{21})$ , which can be written as  $2m_j = \text{Tr } M\sigma_j$ , with  $\sigma_0 := 1$ . When the  $m_0$  and  $\vec{m}$  are real, M is hermitian. Matrix products are easily evaluated using (112). As a simple example we find

$$(m_0 \mathbb{1} + \vec{m} \cdot \vec{\sigma})(m_0 \mathbb{1} - \vec{m} \cdot \vec{\sigma}) = m_0^2 - m_x^2 - m_y^2 - m_z^2 = \det M$$

(remember det  $\alpha \mathbb{1}_N = \alpha^N$ ). Thus, when det  $M \neq 0$ , the inverse of M is

$$M = m_0 \mathbb{1} + \vec{m} \cdot \vec{\sigma} \quad \rightsquigarrow \quad M^{-1} = (m_0 \mathbb{1} - \vec{m} \cdot \vec{\sigma}) / \det M.$$
(114)

For a unitary matrix  $U = u_0 + \vec{u} \cdot \vec{\sigma}$  with det U = 1 we then see from  $U^{\dagger} = u_0^* + \vec{u}^* \cdot \vec{\sigma} \stackrel{!}{=} u_0 - \vec{u} \cdot \vec{\sigma}$ that  $u_0$  must be real and  $\vec{u} = i\vec{n}$  imaginary, so that  $1 = \det U = u_0^2 + ||\vec{n}||^2$ , which allows us to write  $u_0 = \cos \alpha$  and  $\vec{n} = \sin \alpha \hat{n}$  with unit vector  $\hat{n} := \vec{n}/||\vec{n}||$  and  $\alpha \in [0, 2\pi)$ . Thus any special unitary  $2 \times 2$  matrix  $U \in SU(2)$  can be written, using  $(\hat{n} \cdot \vec{\sigma})^2 = 1$  from (112) in the power series,

$$U_{\hat{n},\alpha} = \cos \alpha \, \mathbb{1} + \mathrm{i} \sin \alpha \, (\hat{n} \cdot \vec{\sigma}) = \exp(\mathrm{i}\alpha \, \hat{n} \cdot \vec{\sigma}). \tag{115}$$

General unitary matrices with det  $U = e^{i\alpha_0}$  have the form  $U = e^{i\alpha_0/2}U_{\hat{n},\alpha}$ .

The U are related to rotations of vectors  $\vec{a} \in \mathbb{R}^3$  via  $U(\vec{a} \cdot \vec{\sigma})U^{\dagger}$ . To see this we remember that  $\vec{a} \cdot \vec{\sigma}$  is a hermitian  $2 \times 2$  matrix with zero trace. By the cyclic property of the trace  $\text{Tr } U(\vec{a} \cdot \vec{\sigma})U^{\dagger} = \text{Tr } \vec{a} \cdot \vec{\sigma}$ , so that there exists a unique  $\vec{a}_U$  with  $U(\vec{a} \cdot \vec{\sigma})U^{\dagger} = \vec{a}_U \cdot \vec{\sigma}$ . This mapping  $\vec{a} \to \vec{a}_U$  is linear,  $U((c\vec{a} + \vec{b}) \cdot \vec{\sigma})U^{\dagger} = cU(\vec{a} \cdot \vec{\sigma})U^{\dagger} + U(\vec{b} \cdot \vec{\sigma})U^{\dagger}$ , and preserves the inner product

$$\vec{a} \cdot \vec{b} = \frac{1}{2} \operatorname{Tr}(\vec{a} \cdot \vec{\sigma})(\vec{b} \cdot \vec{\sigma}) = \frac{1}{2} \operatorname{Tr} U(\vec{a} \cdot \vec{\sigma})U^{\dagger} U(\vec{b} \cdot \vec{\sigma})U^{\dagger} = \frac{1}{2} \operatorname{Tr}(\vec{a}_{U} \cdot \vec{\sigma})(\vec{b}_{U} \cdot \vec{\sigma}) = \vec{a}_{U} \cdot \vec{b}_{U}$$

so that it must be a proper rotation,  $\vec{a}_U = R_U \vec{a}$  with  $R_U \in SO(3)$ . To identify which rotation, we consider the special case  $\vec{a}_{\parallel} = c_a \hat{n}$  for which, by (112),  $\vec{a}_{\parallel} \cdot \vec{\sigma}$  commutes with  $\hat{n} \cdot \vec{\sigma}$  so that  $U(\vec{a}_{\parallel} \cdot \vec{\sigma})U^{\dagger} = \vec{a}_{\parallel} \cdot \vec{\sigma}$ , i.e.,  $\hat{n}$  is the axis of rotation. To find the rotation angle  $\vartheta$ , we consider a unit vector  $\hat{a}_{\perp}$  perpendicular to  $\hat{n}$ , for which, using  $(\hat{a}_{\perp} \cdot \vec{\sigma})(\hat{n} \cdot \vec{\sigma}) = i(\hat{a}_{\perp} \times \hat{n}) \cdot \vec{\sigma}$  and  $\operatorname{Tr} \vec{v} \cdot \vec{\sigma} = 0$ ,

$$\cos \vartheta = \hat{a}_{\perp} \cdot R_U \hat{a}_{\perp} = \frac{1}{2} \operatorname{Tr}(\hat{a}_{\perp} \cdot \vec{\sigma}) U(\hat{a}_{\perp} \cdot \vec{\sigma}) U^{\dagger}$$
  
$$= \frac{1}{2} \operatorname{Tr}(\hat{a}_{\perp} \cdot \vec{\sigma})(\cos \alpha + i \sin \alpha (\hat{n} \cdot \vec{\sigma})) \quad (\hat{a}_{\perp} \cdot \vec{\sigma})(\cos \alpha - i \sin \alpha (\hat{n} \cdot \vec{\sigma}))$$
  
$$= \frac{1}{2} \operatorname{Tr}(\cos \alpha \hat{a}_{\perp} - \sin \alpha (\hat{a}_{\perp} \times \hat{n})) \cdot \vec{\sigma} \quad (\cos \alpha \hat{a}_{\perp} + \sin \alpha (\hat{a}_{\perp} \times \hat{n})) \cdot \vec{\sigma}$$
  
$$= (\cos \alpha \hat{a}_{\perp} - \sin \alpha (\hat{a}_{\perp} \times \hat{n})) \cdot (\cos \alpha \hat{a}_{\perp} + \sin \alpha (\hat{a}_{\perp} \times \hat{n})) = (\cos \alpha)^2 - (\sin \alpha)^2 = \cos 2\alpha$$

Hence,  $U_{\hat{n},\alpha} \in \mathrm{SU}(2)$  induces a rotation  $R_U \in \mathrm{SO}(3)$  about the axis  $\hat{n}$  through the angle  $\vartheta = 2\alpha$ . Therefore, matrices in SU(2) are commonly written using the angle of rotation  $\vartheta \in [0, 4\pi)$  instead of  $\alpha \in [0, 2\pi)$  as  $U(\hat{n}, \vartheta) = \exp(\mathrm{i}\frac{\vartheta}{2}\,\hat{n}\cdot\vec{\sigma})$ . We see, in particular, that the two matrices  $U(\hat{n}, \vartheta)$  and  $U(\hat{n}, \vartheta + 2\pi) = -U(\hat{n}, \vartheta)$  in SU(2) induce the same rotation  $R_{-U} = R_U \in \mathrm{SO}(3)$ .

Diagonalizing a hermitian  $2 \times 2$  matrix  $M = m_0 \mathbb{1} + \vec{m} \cdot \vec{\sigma}$  is now simple: just rotate  $\hat{m} \rightarrow \hat{z}$ 

$$U(m_0 \mathbb{1} + \|\vec{m}\| (\hat{m} \cdot \vec{\sigma}))U^{\dagger} = m_0 \mathbb{1} + \|\vec{m}\| \sigma_z$$

from which we easily read off the eigenvalues

$$\varepsilon_{\pm} = m_0 \pm \|\vec{m}\| = \frac{m_{11} + m_{22}}{2} \pm \sqrt{\frac{(m_{11} - m_{22})^2}{4}} + |m_{12}|^2 = \frac{1}{2} \operatorname{Tr} M \pm \sqrt{\left(\frac{1}{2} \operatorname{Tr} M\right)^2 - \det M},$$
  
while the eigenvalues are the columns vectors of  $U^{\dagger} = (\boldsymbol{v}_+, \, \boldsymbol{v}_-)$ 

$$m_0 \mathbb{1} + \vec{m} \cdot \vec{\sigma} = U^{\dagger} (m_0 \mathbb{1} + \|\vec{m}\| \sigma_z) U = m_0 \mathbb{1} + \|\vec{m}\| (\boldsymbol{v}_+, \boldsymbol{v}_-) \sigma_z \begin{pmatrix} \boldsymbol{v}_+^{\dagger} \\ \boldsymbol{v}_-^{\dagger} \end{pmatrix}$$

We still need to determine a U that rotates  $\hat{m} \to \hat{z}$ . The rotation axis should be orthogonal to both vectors, i.e.,  $\hat{n} = \hat{z} \times \hat{m}/\|\hat{z} \times \hat{m}\| = (m_x \hat{y} - m_y \hat{x})/\sqrt{m_x^2 + m_y^2}$ , so that the rotation angle  $\vartheta \in [0, \pi]$  is determined by  $\cos \vartheta = \hat{m} \cdot \hat{z} = m_z/\|\vec{m}\|$ . Using also the other spherical coordinates  $m_x = \|\vec{m}\| \sin \vartheta \cos \varphi$  and  $m_y = \|\vec{m}\| \sin \vartheta \sin \varphi$ , we get  $\hat{n} = \cos \varphi \, \hat{y} - \sin \varphi \, \hat{x}$  so that

$$U(\hat{n},\vartheta) = \begin{pmatrix} \cos\frac{\vartheta}{2} + in_z \sin\frac{\vartheta}{2} & (n_y + in_x) \sin\frac{\vartheta}{2} \\ -(n_y - in_x) \sin\frac{\vartheta}{2} & \cos\frac{\vartheta}{2} - in_z \sin\frac{\vartheta}{2} \end{pmatrix} = \begin{pmatrix} \cos\frac{\vartheta}{2} & +e^{-i\varphi} \sin\frac{\vartheta}{2} \\ -e^{+i\varphi} \sin\frac{\vartheta}{2} & \cos\frac{\vartheta}{2} \end{pmatrix}$$

from which we read off the eigenvectors as the columns of  $U^{\dagger}$  (which you may want to check for simple cases like  $M = \sigma_z$ ,  $\sigma_x$  or  $\sigma_y$ )

$$\boldsymbol{v}_{+} = \begin{pmatrix} \cos\frac{\vartheta}{2} \\ +e^{+i\varphi}\sin\frac{\vartheta}{2} \end{pmatrix} \text{ and } \boldsymbol{v}_{-} = \begin{pmatrix} -e^{-i\varphi}\sin\frac{\vartheta}{2} \\ \cos\frac{\vartheta}{2} \end{pmatrix} \text{ with } \begin{array}{c} \varphi = \arg(m_{21}) = -\arg(m_{12}) \\ \vartheta = \arccos\frac{m_{11} - m_{22}}{\varepsilon_{+} - \varepsilon_{-}} \end{array}$$
(116)

A more symmetric form of the eigenvectors may be obtained by writing  $e^{\mp i\varphi/2} v_{\pm}$ .

## A.3 Some useful commutation relations

Commuting an operator through a product of operators is straightforward

$$AB_{1} \cdots B_{N} = [A, B_{1}] B_{2} \cdots B_{N} + B_{1}AB_{2} \cdots B_{N}$$
  
=  $[A, B_{1}] B_{2} \cdots B_{N} + B_{1}[A, B_{2}] \cdots B_{N} + B_{1}B_{2}A \cdots B_{N}$   
 $\vdots$   
=  $\sum_{n=1}^{N} \prod_{i=1}^{n-1} B_{i} [A, B_{n}] \prod_{i=n+1}^{N} B_{i} + B_{1} \cdots B_{N}A$ 

while, working analogously, anticommuting introduces alternating signs

$$AB_{1} \cdots B_{N} = \{A, B_{1}\} B_{2} \cdots B_{N} - B_{1}AB_{2} \cdots B_{N}$$
  
=  $\{A, B_{1}\} B_{2} \cdots B_{N} - B_{1}\{A, B_{2}\} \cdots B_{N} - B_{1}B_{2}A \cdots B_{N}$   
:  
$$\vdots$$
  
=  $\sum_{n=1}^{N} (-1)^{n-1} \prod_{i=1}^{n-1} B_{i} \{A, B_{n}\} \prod_{i=n+1}^{N} B_{i} + (-1)^{N}B_{1} \cdots B_{N}A$ 

The following special cases are particularly useful

$$[A, BC] = [A, B]C + B[A, C]$$
  
= {A, B}C - B{A, C}  
$$[AB, C] = A[B, C] + [A, C]B$$
  
= A{B, C} - {A, C}B  
$$[AB, CD] = A[B, C]D + AC[B, D] + [A, C]DB + C[A, D]B$$
  
= A{B, C}D - AC{B, D} + {A, C}DB - C{A, D}B

Important examples are  $[c_i^{\dagger}c_j, c_{\gamma}^{\dagger}] = \langle j|\gamma \rangle c_i^{\dagger}$  and  $[c_i^{\dagger}c_j, c_{\gamma}] = -\langle i|\gamma \rangle c_j$ . For the commutator of products of creation and annihilation operators appearing in one- and two-body operators we find

$$\left[c_{i}^{\dagger}c_{j}, c_{\alpha}^{\dagger}c_{\beta}\right] = \left[c_{i}^{\dagger}c_{j}, c_{\alpha}^{\dagger}\right]c_{\beta} + c_{\alpha}^{\dagger}\left[c_{i}^{\dagger}c_{j}, c_{\beta}\right] = \langle j|\alpha\rangle c_{i}^{\dagger}c_{\beta} - \langle \beta|i\rangle c_{\alpha}^{\dagger}c_{j}$$

and

$$\left[c_{i}^{\dagger}c_{j}^{\dagger}c_{k}c_{l}, c_{\alpha}^{\dagger}c_{\beta}\right] = \langle l|\alpha\rangle c_{i}^{\dagger}c_{j}^{\dagger}c_{k}c_{\beta} + \langle k|\alpha\rangle c_{i}^{\dagger}c_{j}^{\dagger}c_{\beta}c_{l} - \langle \beta|j\rangle c_{i}^{\dagger}c_{\alpha}^{\dagger}c_{k}c_{l} - \langle \beta|i\rangle c_{\alpha}^{\dagger}c_{j}^{\dagger}c_{k}c_{l}.$$

# References

- [1] P. Forrest: *The Identity of Indiscernibles*, in The Stanford Encyclopedia of Philosophy http://plato.stanford.edu/entries/identity-indiscernible/
- [2] W. Pauli: The connection between spin and statistics, Phys. Rev. 58, 716 (1940)
- [3] R.F. Streater and A.S. Wightman: *PCT, Spin and Statistics, and All That* (Benjamin, New York, 1964)
- [4] P.-O. Löwdin: Quantum Theory of Many-Particle Systems I, Phys. Rev. 97, 1474 (1955)
- [5] P. Jordan and O. Klein: *Zum Mehrkörperproblem in der Quantenmechanik*, Z. Physik 45, 751 (1927)
- [6] P. Jordan and E. Wigner: Über das Paulische Aquivalenzverbot, Z. Physik 47, 631 (1928)
- [7] R. Peierls: Surprises in Theoretical Physics (Princeton University Press, 1979)
- [8] P. Jorgensen and J. Simons: Second Quantization-Based Methods in Quantum Chemistry (Academic Press, New York, 1981)
- [9] D.J. Thouless: *Stability conditions and nuclear rotations in the Hartree-Fock theory*, Nucl. Phys. 21, 225 (1960)
- [10] Le Thi Hoai: Product wave-functions in Fock-space (MSc Thesis, German Research School for Simulation Sciences, 2015)
- [11] E. Koch: *Many-Electron States* in [17]
- [12] S. Zhang and D.M. Ceperley: Hartree-Fock Ground State of the Three-Dimensional Electron Gas, Phys. Rev. Lett. 100, 236404 (2008)
- [13] E. Koch: *Exchange Mechanisms* in [18]
- [14] A. Szabo and N.S. Ostlund: Modern Quantum Chemistry (McGraw Hill, New York, 1989)
- [15] P.-O. Löwdin: On the Non-Orthogonality Problem Connected with the Use of Atomic Wave Functions in the Theory of Molecules and Crystals, J. Chem. Phys. 18, 365 (1950)
- [16] W. Pauli jr.: Zur Quantenmechanik des magnetischen Elektrons, Z. Phys. 43, 601 (1927)
- [17] E. Pavarini, E. Koch, and U. Schollwöck (Eds.): *Emergent Phenomena in Correlated Matter*, Reihe Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013) http://www.cond-mat.de/events/correl13
- [18] E. Pavarini, E. Koch, F. Anders, and M. Jarrell (Eds.): From Models to Materials, Reihe Modeling and Simulation, Vol. 2 (Forschungszentrum Jülich, 2012) http://www.cond-mat.de/events/correll2