1 Second Quantization and Jordan-Wigner Representations

Erik Koch
Jülich Supercomputer Centre
Forschungszentrum Jülich

Contents

1 Many-electron states **2**

2 Second quantization **5**
2.1 Creation and annihilation operators 5
2.2 Representation of Slater determinants 7
2.3 Representation of n-body operators 8
2.4 Transforming the orbital basis 12

3 Exact diagonalization **13**
3.1 Variational principles 13
3.2 Matrix eigenvalue problem 15
3.3 Dimension of the Hilbert space and sparseness 16
3.4 Non-interacting electrons 18

4 Jordan-Wigner representation **18**
4.1 Spins, hard-core bosons, and fermions 22

A Appendix **24**
A.1 Atomic units 24
A.2 Non-orthonormal basis 25
A.3 Pauli matrices 26
A.4 Some useful commutation relations 28

E. Pavarini and E. Koch (eds.)
Simulating Correlations with Computers
Modeling and Simulation Vol. 11
Forschungszentrum Jülich, 2021, ISBN 978-3-95806-529-1
http://www.cond-mat.de/events/correl21

1 Many-electron states

One of the great mysteries 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. Consequently, every observable for, e.g., electrons, must treat each electron in the same way. Thus, in principle, observables must act on all the 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 ourselves to the electrons in the system under consideration, assuming that the differential overlap with all other electrons vanishes. Any observable $M(x_1,\dots,x_N)$ for the N electrons in our system must then be symmetric under permutations of the coordinates 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

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

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

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 wave functions $\Psi(x_1,x_2,\cdots)$ the situation is slightly more complex. Since the probability density $|\Psi(x_1,x_2,\cdots)|^2$ is an observable, the wave function 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 wave functions 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 wave function we thus obtain a many-electron wavefunction by antisymmetrizing

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

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 φ_α can be antisymmetrized much more efficiently (in $\mathcal{O}(N^3)$ steps) by writing it in the form of a determinant

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

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 $\Phi_{\alpha,\alpha'}(x,x') = \big(\varphi_\alpha(x)\varphi_{\alpha'}(x') - \varphi_{\alpha'}(x)\varphi_\alpha(x')\big)/\sqrt{2}$ 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

$$\sum_n \overline{\varphi_{\alpha_n}(x)}\,\varphi_{\alpha_n}(x') = \delta(x{-}x') \text{ (complete)} \qquad \int dx\, \overline{\varphi_{\alpha_n}(x)}\,\varphi_{\alpha_m}(x) = \delta_{n,m} \text{ (orthonormal)}\,. \tag{5}$$

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 the $x_2,\ldots,x_N$ kept fixed. We can then expand it in the complete set $\{\varphi_{\alpha_n}\}$ as

$$a(x_1,\ldots,x_N) = \sum_{\alpha_1} a_{\alpha_1}(x_2,\ldots,x_N)\,\varphi_{\alpha_1}(x_1)$$

with expansion coefficients that depend on the remaining coordinates

$$a_{\alpha_1}(x_2,\ldots,x_N) = \int dx_1\, \overline{\varphi_{\alpha_1}(x_1)}\,a(x_1,x_2,\ldots,x_N)\,.$$

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

$$a_{\alpha_1}(x_2,\ldots,x_N) = \sum_{\alpha_2} a_{\alpha_1,\alpha_2}(x_3,\ldots,x_N)\,\varphi_{\alpha_2}(x_2)\,.$$

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

$$a(x_1,\ldots,x_N) = \sum_{\alpha_1,\ldots,\alpha_N} a_{\alpha_1,\ldots,\alpha_N}\,\varphi_{\alpha_1}(x_1)\cdots\varphi_{\alpha_N}(x_N)\,.$$

When the N-particle function is antisymmetric, the expansion coefficients must be antisymmetric under permutation of the indices: $a_{\alpha_{p(1)},\ldots,\alpha_{p(N)}} = (-1)^P a_{\alpha_1,\ldots,\alpha_N}$. Fixing some particular order of the indices, e.g., $\alpha_1 < \alpha_2 < \ldots < \alpha_N$, we thus get an expansion in Slater determinants

$$\Psi(x_1,\ldots,x_N) = \sum_{\alpha_1<\ldots<\alpha_N} a_{\alpha_1,\ldots,\alpha_N} \underbrace{\sum_P (-1)^P \varphi_{\alpha_{p(1)}}(x_1)\cdots\varphi_{\alpha_{p(N)}}(x_N)}_{=\sqrt{N!}\,\Phi_{\alpha_1,\ldots,\alpha_N}(x_1,\ldots,x_N)}\,.$$

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

$$\Big\{\Phi_{\alpha_1,\ldots,\alpha_N}(x_1,\ldots,x_N)\,\Big|\;\alpha_1<\alpha_2<\cdots<\alpha_N\Big\} \tag{6}$$

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

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

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 $\Phi_{\alpha_1,\dots,\alpha_N}(\boldsymbol{x})$, this representation is not unique: Any unitary transformation among the N occupied orbitals will not change the determinant (up to a phase). It is thus sufficient to give the N-dimensional subspace spanned by the orbitals $\varphi_1, \dots, \varphi_N$ in the single-electron Hilbert space. The projector to this space is the one-body density matrix

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

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

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

where $\Phi_{\alpha_{i\neq n}}(x_2,\dots,x_N)$ 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 α_n removed. Inserting this into (8) we find

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

which is the expansion of the one-body density matrix in eigenfunctions (natural orbitals), with eigenvalues (natural occupation numbers) either one or zero. Due to the degeneracy of the natural occupation numbers the natural orbitals are not unique. Any many-electron wave function $\Psi(\boldsymbol{x})$ with the same one-body density matrix $\Gamma^{(1)}_{\Phi}$ 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, 5]. 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 wave functions. It is called second quantization.

2 Second quantization

The idea behind the second quantized approach to writing many-body wave functions is remarkably simple: When writing Slater determinants in the form (4) we are working in a real-space basis. It is, however, often simpler to consider abstract states: Instead of a wave function $\varphi_\alpha(x)$, we write a Dirac state $|\alpha\rangle$. Second quantization allows us to do the same for Slater determinants. Let us consider a Slater determinant for two electrons, one in state $\varphi_\alpha(x)$, the other in state $\varphi_\beta(x)$. It is simply the antisymmetrized product of the two states

$$\Phi_{\alpha\beta}(x_1, x_2) = \frac{1}{\sqrt{2}} \left(\varphi_\alpha(x_1)\varphi_\beta(x_2) - \varphi_\beta(x_1)\varphi_\alpha(x_2)\right). \tag{11}$$

This expression is quite cumbersome because we explicitly specify the coordinates. We can try to get rid of the coordinates by defining a two-particle Dirac state

$$|\alpha, \beta\rangle := \frac{1}{\sqrt{2}} \left(|\alpha\rangle|\beta\rangle - |\beta\rangle|\alpha\rangle\right).$$

While the expression is already simpler, we still have to keep track of the order of the particles by specifying the position of the kets. The idea of second quantization is to specify the states using operators

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle. \tag{12}$$

Now the order of the particles is specified by the order of the operators. To ensure the antisymmetry of the wave function the operators have to change sign when they are reordered

$$|\alpha, \beta\rangle = c_\beta^\dagger c_\alpha^\dagger |0\rangle = -c_\alpha^\dagger c_\beta^\dagger |0\rangle = -|\beta, \alpha\rangle\,. \tag{13}$$

2.1 Creation and annihilation operators

To arrive at the formalism of second quantization we postulate a set of operators that have certain reasonable properties. We then verify that we can use these operators to represent Slater determinants. But first we consider a few simple states to motivate what properties the new operators ought to have.

To be able to construct many-electron states, we start from the simplest such state: $|0\rangle$ the *vacuum state* with no electron, which we assume to be normalized $\langle 0|0\rangle = 1$. Next we introduce for each single-electron state $|\alpha\rangle$ an operator $c_\alpha^\dagger$ such that $c_\alpha^\dagger|0\rangle = |\alpha\rangle$. We call them *creation operators* since they add an electron (in state α) to the state that they act on: in $c_\alpha^\dagger|0\rangle$ the creation operator adds an electron to the vacuum state (N=0), resulting in a single-electron state (N=1). Applying another creation operator produces a two-electron state $c_\beta^\dagger c_\alpha^\dagger|0\rangle$, ($N$=2). To ensure the antisymmetry of the two electron state, the product of creation operators has to change sign when they are reordered: $c_\alpha^\dagger c_\beta^\dagger = -c_\beta^\dagger c_\alpha^\dagger$. This is more conveniently written as $\{c_\alpha^\dagger,\, c_\beta^\dagger\} = 0$ by introducing the *anti-commutator*

$$\{A,\, B\} := A\,B + B\,A\,. \tag{14}$$

As we have seen, the simplest state we can produce with the creation operators is the single-electron state $|\alpha\rangle = c_\alpha^\dagger|0\rangle$. When we want to calculate its norm, we have to consider the adjoint of $c_\alpha^\dagger|0\rangle$, formally obtaining $\langle\alpha|\alpha\rangle = \langle 0|c_\alpha c_\alpha^\dagger|0\rangle$, or, more generally, $\langle\alpha|\beta\rangle = \langle 0|c_\alpha c_\beta^\dagger|0\rangle$. This implies that c_α, the adjoint of a creation operator, must remove an electron from the state, otherwise the overlap of $c_\alpha c_\beta^\dagger|0\rangle$ with the vacuum state $\langle 0|$ would vanish. We therefore call the adjoint of the creation operator an *annihilation operator*. We certainly cannot take an electron out of the vacuum state, so $c_\alpha|0\rangle = 0$. To obtain the overlap of one-electron states we postulate the anticommutation relation $\{c_\alpha, c_\beta^\dagger\} = \langle\alpha|\beta\rangle$, giving $\langle 0|c_\alpha c_\beta^\dagger|0\rangle = \langle 0|\{c_\alpha,\, c_\beta^\dagger\} - c_\beta^\dagger c_\alpha|0\rangle = \langle\alpha|\beta\rangle$. For completeness, taking the adjoint of the anticommutation relation for the creation operators, we obtain the corresponding anticommutator of the annihilators: $\{c_\alpha,\, c_\beta\} = 0$.
Thus, we are led to define the vacuum state $|0\rangle$ and the set of operators c_α related to single-electron states $|\alpha\rangle$ with the properties

$$\boxed{\begin{aligned} c_\alpha|0\rangle &= 0 \qquad & \{c_\alpha,\, c_\beta\} &= 0 = \{c_\alpha^\dagger,\, c_\beta^\dagger\} \\ \langle 0|0\rangle &= 1 \qquad & \{c_\alpha,\, c_\beta^\dagger\} &= \langle\alpha|\beta\rangle \end{aligned}} \tag{15}$$

As a direct consequence we obtain the Pauli principle in the form $c_\alpha c_\alpha = 0 = c_\alpha^\dagger c_\alpha^\dagger$.
We note that the creators transform in the same way as the single-electron states they represent

$$|\tilde\alpha_i\rangle = \sum_\mu |\alpha_\mu\rangle\, U_{\mu i} \qquad \rightsquigarrow \quad \tilde c_{\tilde\alpha_i}^\dagger|0\rangle = \sum_\mu c_{\alpha_\mu}^\dagger|0\rangle\, U_{\mu i} = \left(\sum_\mu c_{\alpha_\mu}^\dagger U_{\mu i}\right)|0\rangle. \tag{16}$$

The creators and annihilators are clearly not operators in a Hilbert space, but transfer states from an N-electron to a $N\pm1$-electron Hilbert space, i.e., they are operators defined on *Fock space*.
It is also remarkable that the mixed anti-commutator is the only place where the orbitals that distinguish different operators enter.
To make contact with the notation of first quantization, we introduce the *field operators* $\hat\Psi^\dagger(x)$, with $x = (r,\sigma)$, that create an electron of spin σ at position r, i.e., in state $|x\rangle = |r,\sigma\rangle$. Given a complete, orthonormal set of orbitals $\{\varphi_n\}$, we can expand $|x\rangle$

$$\hat\Psi^\dagger(x)|0\rangle = |x\rangle = \sum_n |\varphi_n\rangle\langle\varphi_n|x\rangle = \sum_n c_{\varphi_n}^\dagger|0\rangle\langle\varphi_n|x\rangle \tag{17}$$

from which we obtain

$$\hat\Psi^\dagger(x) = \sum_n \overline{\langle x|\varphi_n\rangle}\, c_{\varphi_n}^\dagger = \sum_n \overline{\varphi_n(x)}\, c_{\varphi_n}^\dagger. \tag{18}$$

The anticommutators then follow from (15) for an orthonormal and complete set, e.g.,

$$\{\hat\Psi(x),\, \hat\Psi^\dagger(x')\} = \sum_{n,m}\langle x|\varphi_n\rangle \underbrace{\{c_{\varphi_n},\, c_{\varphi_m}^\dagger\}}_{=\delta_{n,m}}\langle\varphi_m|x'\rangle = \sum_n \langle x|\varphi_n\rangle\langle\varphi_n|x'\rangle = \langle x|x'\rangle = \delta(x-x'),$$

resulting in the anticommutation relations for the field operators

$$\{\hat\Psi(x),\, \hat\Psi(x')\} = 0 = \{\hat\Psi^\dagger(x),\, \hat\Psi^\dagger(x')\} \quad \text{and} \quad \{\hat\Psi(x),\, \hat\Psi^\dagger(x')\} = \langle x|x'\rangle. \tag{19}$$

We can, of course, expand the field operators also in a non-orthogonal set of orbitals $\{|\chi_i\rangle\}$, as long as it is complete, $\sum_{i,j} |\chi_i\rangle (S^{-1})_{ij} \langle\chi_j| = \mathbb{1}$, where $S_{ij} = \langle\chi_i|\chi_j\rangle$ is the overlap matrix

$$\hat{\Psi}^\dagger(x) = \sum_{i,j} c_i^\dagger \, (S^{-1})_{ij} \, \langle\chi_j|x\rangle. \tag{20}$$

Conversely, given any single-electron wave functions in real space $\varphi(x)$, we can express the corresponding creation operator in terms of the field operators

$$c_\varphi^\dagger = \int dx \, \varphi(x) \, \hat{\Psi}^\dagger(x). \tag{21}$$

Its anticommutator with the field annihilator just gives back the single-electron wave function

$$\left\{\hat{\Psi}(x), \, c_\varphi^\dagger\right\} = \int dx' \, \varphi(x') \left\{\hat{\Psi}(x), \, \hat{\Psi}^\dagger(x')\right\} = \varphi(x) \, . \tag{22}$$

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:

$$\Phi_{\alpha_1\alpha_2\ldots\alpha_N}(x_1, x_2, \ldots, x_N) = \frac{1}{\sqrt{N!}} \left\langle 0\right| \hat{\Psi}(x_1)\hat{\Psi}(x_2)\ldots\hat{\Psi}(x_N) \;\; c_{\alpha_N}^\dagger \ldots c_{\alpha_2}^\dagger c_{\alpha_1}^\dagger \left|0\right\rangle. \tag{23}$$

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 (23) is by induction. As a warm-up we consider the case of a single-electron wave function (N=1). Using the anticommutation relation (22), we see that

$$\left\langle 0\right| \hat{\Psi}(x_1) \, c_{\alpha_1}^\dagger \left|0\right\rangle = \left\langle 0\right| \varphi_{\alpha_1}(x_1) - c_{\alpha_1}^\dagger \hat{\Psi}(x_1) \left|0\right\rangle = \varphi_{\alpha_1}(x_1) \, . \tag{24}$$

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

$$\begin{aligned} \left\langle 0\right| \hat{\Psi}(x_1)\hat{\Psi}(x_2) \, c_{\alpha_2}^\dagger c_{\alpha_1}^\dagger \left|0\right\rangle &= \left\langle 0\right| \hat{\Psi}(x_1) \big(\varphi_{\alpha_2}(x_2) - c_{\alpha_2}^\dagger \hat{\Psi}(x_2)\big) c_{\alpha_1}^\dagger \left|0\right\rangle \\ &= \left\langle 0\right| \hat{\Psi}(x_1) c_{\alpha_1}^\dagger \left|0\right\rangle \varphi_{\alpha_2}(x_2) - \left\langle 0\right| \hat{\Psi}(x_1) c_{\alpha_2}^\dagger \hat{\Psi}(x_2) c_{\alpha_1}^\dagger \left|0\right\rangle \\ &= \varphi_{\alpha_1}(x_1)\varphi_{\alpha_2}(x_2) - \varphi_{\alpha_2}(x_1)\varphi_{\alpha_1}(x_2) \, . \end{aligned} \tag{25}$$

We see how anticommuting automatically produces the appropriate signs for the antisymmetric wave function. 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 terms with alternating sign

$$\begin{aligned} &\left\langle 0\right| \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_{N-1})\hat{\Psi}(x_N) \, c_{\alpha_N}^\dagger c_{\alpha_{N-1}}^\dagger \cdots c_{\alpha_1}^\dagger \left|0\right\rangle \; = \\ &\qquad\qquad\begin{array}{lll} + & \left\langle 0\right| \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_{N-1}) \, c_{\alpha_{N-1}}^\dagger \cdots c_{\alpha_1}^\dagger \left|0\right\rangle & \varphi_{\alpha_N}(x_N) \\ - & \left\langle 0\right| \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_{N-1}) \, \prod_{n\neq N-1} c_{\alpha_n}^\dagger \left|0\right\rangle & \varphi_{\alpha_{N-1}}(x_N) \\ & \vdots & \\ (-1)^{N-1} & \left\langle 0\right| \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_{N-1}) \, c_{\alpha_N}^\dagger \;\; \cdots \;\; c_{\alpha_2}^\dagger \left|0\right\rangle & \varphi_{\alpha_1}(x_N) \, . \end{array} \end{aligned}$$

Using (23) 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 (23) 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 wave functions—just as in Dirac notation. In particular we can rewrite the basis of Slater determinants (6) into a basis of product states

$$\left\{ c^\dagger_{\alpha_N} \cdots c^\dagger_{\alpha_1} |0\rangle \;\middle|\; \alpha_1 < \cdots < \alpha_N \right\}, \tag{26}$$

which allows us to express any N-electron state as

$$|\Psi\rangle = \sum_{\alpha_1<\cdots<\alpha_N} a_{\alpha_1,\ldots,\alpha_N}\, c^\dagger_{\alpha_N} \cdots c^\dagger_{\alpha_1} |0\rangle. \tag{27}$$

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 $M(\boldsymbol{x})$ 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:

$$\int dx\, \overline{\varphi_n(x)}\, M(x)\, \varphi_m(x) = \langle \varphi_n | \underbrace{\int dx\, |x\rangle M(x) \langle x|}_{=:\hat{M}} \varphi_m \rangle = \langle \varphi_n | \hat{M} | \varphi_m \rangle. \tag{28}$$

For N-electron Slater determinants this becomes

$$\int dx_1 \cdots dx_N\, \overline{\Phi_{\beta_1\cdots\beta_N}(x_1,\cdots,x_N)}\, M(x_1,\ldots,x_N)\, \Phi_{\alpha_1\cdots\alpha_N}(x_1,\cdots,x_N)$$

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

$$= \left\langle 0 \right| c_{\beta_1} \cdots c_{\beta_N}\, \hat{M}\, c^\dagger_{\alpha_N} \cdots c^\dagger_{\alpha_1} \left| 0 \right\rangle$$

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

$$\hat{M} = \frac{1}{N!} \int dx_1 \cdots x_N\, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1)\, M(x_1,\cdots,x_N)\, \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N)\,. \tag{29}$$

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 $|0\rangle\langle 0| = \mathbb{1}_0$. Keeping this in mind, we can work entirely in terms of our algebra (15).

To see what (29) means we look, in turn, at the different n-body parts of $M(\boldsymbol{x})$, (2):

$$M(\mathbf{x}) = M^{(0)} + \sum_i M^{(1)}(x_i) + \sum_{i<j} M^{(2)}(x_i, x_j) + \sum_{i<j<k} M^{(3)}(x_i, x_j, x_k) + \cdots \tag{30}$$

We start with the simplest case, the zero-body operator, which, up to a trivial prefactor, is $M^{(0)}(x_1, \cdots, x_N) = 1$. Operating on an N-electron state, it gives

$$\begin{aligned}
\hat{M}^{(0)} &= \frac{1}{N!} \int dx_1 dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \hat{\Psi}^\dagger(x_1) \, \hat{\Psi}(x_1) \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \int \quad dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \qquad \hat{N} \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \int \quad dx_2 \cdots x_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_2) \qquad 1 \qquad \hat{\Psi}(x_2) \cdots \hat{\Psi}(x_N) \\
&\;\;\vdots \\
&= \frac{1}{N!} \, 1 \cdot 2 \, \cdots \, N = 1 \,,
\end{aligned} \tag{31}$$

where we have used that the operator

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

counts the number of electrons: Applied to the vacuum state it gives $\hat{N} \, |0\rangle = 0$, while each creation operator increases the count by one: $\hat{N} c_n^\dagger = c_n^\dagger (\hat{N}{+}1)$, which follows from

$$[\hat{N}, \, c_n^\dagger] = \int dx \, [\hat{\Psi}^\dagger(x) \hat{\Psi}(x), \, c_n^\dagger] = \int dx \, \hat{\Psi}^\dagger(x) \, \{\hat{\Psi}(x), \, c_n^\dagger\} = \int dx \, \hat{\Psi}^\dagger(x) \, \varphi_n(x) = c_n^\dagger, \tag{32}$$

where we have used the simple relation $[AB, \, C] = A\{B, \, C\} - \{A, \, C\}B$. Adjoining this relation gives a minus sign, i.e., annihilators reduce the count by one: $\hat{N} c_n = c_n (\hat{N}{-}1)$. 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 (31) just means that the overlap of two Slater determinants (7) is equal to that of the corresponding product states

$$\int d\boldsymbol{x} \, \overline{\Phi_{\alpha_1, \ldots, \alpha_N}(\boldsymbol{x})} \, \Phi_{\beta_1, \ldots, \beta_N}(\boldsymbol{x}) = \big\langle 0 \big| \, c_{\alpha_1} \cdots c_{\alpha_N} \, c_{\beta_N}^\dagger \cdots c_{\beta_1}^\dagger \big| 0 \big\rangle. \tag{33}$$

2.3.1 One-body operators

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

$$\begin{aligned}\hat{M}^{(1)} &= \frac{1}{N!}\int dx_1\cdots dx_N\,\hat\Psi^\dagger(x_N)\cdots\hat\Psi^\dagger(x_1)\sum_j M^{(1)}(x_j)\;\hat\Psi(x_1)\cdots\hat\Psi(x_N)\\ &= \frac{1}{N!}\sum_j\int dx_j\,\hat\Psi^\dagger(x_j)\,M^{(1)}(x_j)\,(N{-}1)!\,\hat\Psi(x_j)\\ &= \frac{1}{N}\sum_j\int dx_j\,\hat\Psi^\dagger(x_j)\,M^{(1)}(x_j)\,\hat\Psi(x_j)\\ &= \int dx\quad\hat\Psi^\dagger(x)\;\;M^{(1)}(x)\;\;\hat\Psi(x)\end{aligned}\tag{34}$$

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

$$\hat{M}^{(1)} = \sum_{n,m}\int dx\,\overline{\varphi_n(x)}\,M(x)\,\varphi_m(x)\;c_n^\dagger c_m = \sum_{n,m}\langle\varphi_n|M^{(1)}|\varphi_m\rangle\;c_n^\dagger c_m = \sum_{n,m} c_n^\dagger\,M^{(1)}_{nm}\,c_m.\tag{35}$$

The matrix elements $M^{(1)}_{nm}=\langle\varphi_n|M^{(1)}|\varphi_m\rangle$ transform like a single-electron matrix $\boldsymbol{M}^{(1)}$: From (16) and writing the annihilation operators as a column vector $\boldsymbol{c}$ we see that

$$\hat{M}^{(1)} = \boldsymbol{c}^\dagger\,\boldsymbol{M}^{(1)}\,\boldsymbol{c}\;\; = \boldsymbol{c}^\dagger\boldsymbol{U}^\dagger\,\boldsymbol{U}\boldsymbol{M}^{(1)}\boldsymbol{U}^\dagger\,\boldsymbol{U}\boldsymbol{c} = \tilde{\boldsymbol{c}}^\dagger\,\tilde{\boldsymbol{M}}^{(1)}\,\tilde{\boldsymbol{c}}\,.\tag{36}$$

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 (35) 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

$$\hat\Gamma^{(1)}(x;x') = \hat\Psi^\dagger(x)\hat\Psi(x')\tag{37}$$

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

$$\hat\Gamma^{(1)}(x;x') = \sum_{n,m}\overline{\varphi_n(x)}\,\varphi_m(x')\,c_n^\dagger c_m\,.\tag{38}$$

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:

$$\begin{aligned}
\hat{M}^{(2)} &= \frac{1}{N!} \int dx_1 \cdots dx_N \, \hat{\Psi}^\dagger(x_N) \cdots \hat{\Psi}^\dagger(x_1) \sum_{i<j} M^{(2)}(x_i, x_j) \, \hat{\Psi}(x_1) \cdots \hat{\Psi}(x_N) \\
&= \frac{1}{N!} \sum_{i<j} \int dx_i dx_j \, \hat{\Psi}^\dagger(x_j) \hat{\Psi}^\dagger(x_i) \, M^{(2)}(x_i, x_j) \, (N-2)! \, \hat{\Psi}(x_i) \hat{\Psi}(x_j) \\
&= \frac{1}{N(N-1)} \sum_{i<j} \int dx_i dx_j \, \hat{\Psi}^\dagger(x_j) \hat{\Psi}^\dagger(x_i) \, M^{(2)}(x_i, x_j) \, \hat{\Psi}(x_i) \hat{\Psi}(x_j) \\
&= \frac{1}{2} \int dx \; dx' \; \hat{\Psi}^\dagger(x') \; \hat{\Psi}^\dagger(x) \;\; M^{(2)}(x, x') \;\; \hat{\Psi}(x) \; \hat{\Psi}(x')
\end{aligned}$$

Expanding in an orthonormal basis, we get

$$\begin{aligned}
\hat{M}^{(2)} &= \frac{1}{2} \sum_{n,n',m,m'} \int dx dx' \, \overline{\varphi_{n'}(x')} \, \overline{\varphi_n(x)} \, M^{(2)}(x, x') \, \varphi_m(x) \varphi_{m'}(x') \quad c^\dagger_{n'} c^\dagger_n c_m c_{m'} \\
&= \frac{1}{2} \sum_{n,n',m,m'} \qquad\qquad \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle \qquad\qquad c^\dagger_{n'} c^\dagger_n c_m c_{m'} \qquad (39)
\end{aligned}$$

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. A more natural choice would be to exchange the order in the bra, reflecting that taking the adjoint of the operators changes their order. Either way, $M_{nn',\,mm'} = \langle \varphi_n \varphi_{n'} | M^{(2)} | \varphi_m \varphi_{m'} \rangle$ transforms like a fourth-order tensor: Transforming to a different basis (16) gives

$$\tilde{M}^{(2)}_{\nu\nu',\mu\mu'} = \sum_{n,n',m,m'} U^\dagger_{\nu n} U^\dagger_{\nu' n'} M_{nn',mm'} U_{m\mu} U_{m'\mu'} \, . \qquad (40)$$

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 $c^\dagger_n c^\dagger_n = \{c^\dagger_n, \, c^\dagger_n\}/2 = 0$, and likewise for $M_{nn',\,mm}$.

Note that the representation (39) 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

$$\hat{M}^{(2)} = \sum_{n'>n, \, m'>m} c^\dagger_{n'} c^\dagger_n \underbrace{\left(M^{(2)}_{nn',\,mm'} - M^{(2)}_{n'n,\,mm'} \right)}_{=: \breve{M}^{(2)}_{nn',\,mm'}} c_m c_{m'} \, . \qquad (41)$$

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

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 (30) that was explicitly formulated in an N-electron Hilbert space, the results (31), (35), and (39) 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 (16) that the creators transform in the same way as the orbitals they represent

$$|\beta_i\rangle = U|\alpha_i\rangle = \sum_j |\beta_j\rangle\langle\alpha_j|\alpha_i\rangle = \sum_\mu |\alpha_\mu\rangle \underbrace{\langle\alpha_\mu|U|\alpha_i\rangle}_{=:U_{\mu i}} \quad \rightsquigarrow \quad c^\dagger_{\beta_i} = \sum_\mu c^\dagger_{\alpha_\mu} U_{\mu i}\,, \tag{42}$$

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

$$c^\dagger_{\beta_i} = e^{\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}}\, c^\dagger_{\alpha_\mu}\, e^{-\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}}\,. \tag{43}$$

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

$$e^{\lambda A} B\, e^{-\lambda A} = B + \lambda\,[A,\,B] + \frac{\lambda^2}{2!}\big[A,\,[A,\,B]\big] + \frac{\lambda^3}{3!}\Big[A,\,\big[A,\,[A,\,B]\big]\Big] + \cdots\,, \tag{44}$$

where the expansion coefficients follow by taking the derivatives of the left hand side at $\lambda = 0$, together with the commutator $[c^\dagger_{\alpha_\mu} c_{\alpha_\nu},\, c^\dagger_{\alpha_\kappa}] = c^\dagger_{\alpha_\mu}\,\delta_{\nu,\kappa}$, from which we obtain for the repeated commutators

$$\Big[\sum_{\mu,\nu} M_{\mu\nu} c^\dagger_{\alpha_\mu} c_{\alpha_\nu},\, \sum_\kappa c^\dagger_{\alpha_\kappa} \left(M^n\right)_{\kappa i}\Big] = \sum_{\mu\nu\kappa} c^\dagger_{\alpha_\mu}\, M_{\mu\nu}\,\delta_{\nu,\kappa} \left(M^n\right)_{\kappa i} = \sum_\mu c^\dagger_{\alpha_\mu} \left(M^{n+1}\right)_{\mu i}\,. \tag{45}$$

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

$$e^{\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}} \prod c^\dagger_{\alpha_n} |0\rangle = \prod e^{\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}}\, c^\dagger_{\alpha_n} e^{-\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}} |0\rangle = \prod c^\dagger_{\beta_n} |0\rangle\,. \tag{46}$$

This is, e.g., used when working in the interaction picture. Anticommutators with transformed operators, (42), are simply $\big\{c_{\alpha_j},\, e^{-\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}}\, c^\dagger_{\alpha_i}\, e^{-\boldsymbol{c}^\dagger \boldsymbol{M} \boldsymbol{c}}\big\} = \langle\alpha_j|e^M|\alpha_i\rangle$.
Annihilation operators, being the adjoint of the creators, transform in just the expected way

$$c_{\beta_i} = e^{-\boldsymbol{c}^\dagger \boldsymbol{M}^\dagger \boldsymbol{c}}\, c_{\alpha_\mu}\, e^{\boldsymbol{c}^\dagger \boldsymbol{M}^\dagger \boldsymbol{c}}\,, \tag{47}$$

which means that for unitary transformations, where $\boldsymbol{M}$ 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 Exact diagonalization

We have worked, so far, with complete, i.e., infinite bases. This is, of course, not possible in actual computer simulations, where we have to confine ourselves to finite basis sets. Such calculations on subspaces are based on the variational principle.

3.1 Variational principles

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

$$E[\Psi] = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle}\,. \tag{48}$$

Its variation is

$$E[\Psi+\delta\Psi] = E[\Psi] + \frac{\langle\delta\Psi|H|\Psi\rangle + \langle\Psi|H|\delta\Psi\rangle}{\langle\Psi|\Psi\rangle} - \langle\Psi|H|\Psi\rangle\,\frac{\langle\delta\Psi|\Psi\rangle + \langle\Psi|\delta\Psi\rangle}{\langle\Psi|\Psi\rangle^2} + \mathcal{O}^2. \tag{49}$$

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

$$H|\Psi_n\rangle = E_n|\Psi\rangle\,, \tag{50}$$

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

$$|\Psi\rangle = \sum_n |\Psi_n\rangle\,\langle\Psi_n|\Psi\rangle \tag{51}$$

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

$$\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\sum_{m,n}\langle\Psi|\Psi_m\rangle\langle\Psi_m|H|\Psi_n\rangle\langle\Psi_n|\Psi\rangle}{\sum_{m,n}\langle\Psi|\Psi_m\rangle\langle\Psi_m|\Psi_n\rangle\langle\Psi_n|\Psi\rangle} = \frac{\sum_n E_n\left|\langle\Psi_n|\Psi\rangle\right|^2}{\sum_n\left|\langle\Psi_n|\Psi\rangle\right|^2}\,. \tag{52}$$

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

$$\frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \frac{\sum_n E_n\left|\langle\Psi_n|\Psi\rangle\right|^2}{\sum_n\left|\langle\Psi_n|\Psi\rangle\right|^2} \geq \frac{\sum_n E_0\left|\langle\Psi_n|\Psi\rangle\right|^2}{\sum_n\left|\langle\Psi_n|\Psi\rangle\right|^2} = E_0\,. \tag{53}$$

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

$$\frac{\langle\Psi_{\perp_n}|H|\Psi_{\perp_n}\rangle}{\langle\Psi_{\perp_n}|\Psi_{\perp_n}\rangle} \geq E_n \quad \text{if } \langle\Psi_i|\Psi_{\perp_n}\rangle = 0 \text{ for } i = 0,\ldots,n{-}1. \tag{54}$$

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_α at position $\boldsymbol{R}_\alpha$,

$$H = -\frac{1}{2}\sum_i \Delta_i - \sum_{i,\alpha}\frac{Z_\alpha}{|\boldsymbol{r}_i - \boldsymbol{R}_\alpha|} + \sum_{i<j}\frac{1}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} + \sum_{\alpha<\beta}\frac{Z_\alpha Z_\beta}{|\boldsymbol{R}_\alpha - \boldsymbol{R}_\beta|}\,, \tag{55}$$

the Schrödinger equation is a partial differential equation. In second quantization it becomes a linear-algebra problem: We introduce an orbital basis set $\{\varphi_k \,|\, k\}$, which for simplicity we assume here to be orthonormal, from which we construct an orthonormal basis of N-electron product states, $\{\Phi_{k_1,\dots,k_N} \,|\, k_1<\dots<k_N\}$. To simplify the notation we sort the basis states, e.g., lexicographically in the orbital indices $\boldsymbol{k} = (k_1,\dots,k_N)$ and define the row vector of basis states $|\boldsymbol{\Phi}\rangle := \big(|\Phi_{\boldsymbol{1}}\rangle, |\Phi_{\boldsymbol{2}}\rangle, \dots\big)$. The expansion of a state $|\Psi\rangle$ in this basis can then be written as

$$|\Psi\rangle = \sum_{k_1<\dots<k_N} a_{k_1,\dots,k_N}\,|\Phi_{k_1,\dots,k_N}\rangle = \sum_i a_i\,|\Phi_i\rangle = |\boldsymbol{\Phi}\rangle\,\boldsymbol{a}\,, \tag{56}$$

where $\boldsymbol{a}$ is the vector of expansion coefficients. Likewise we can write the Schrödinger equation as a matrix eigenvalue problem

$$\boldsymbol{H}\boldsymbol{a} = \langle\boldsymbol{\Phi}|\hat{H}|\boldsymbol{\Phi}\rangle\boldsymbol{a} = \begin{pmatrix} \langle\Phi_1|\hat{H}|\Phi_1\rangle & \langle\Phi_1|\hat{H}|\Phi_2\rangle\cdots \\ \langle\Phi_2|\hat{H}|\Phi_1\rangle & \langle\Phi_2|\hat{H}|\Phi_2\rangle\cdots \\ \vdots & \vdots \quad \ddots \end{pmatrix}\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E\begin{pmatrix} a_1 \\ a_2 \\ \vdots \end{pmatrix} = E\boldsymbol{a}\,. \tag{57}$$

From the eigenvectors of the matrix $\boldsymbol{H}$ we easily recover the eigenstates of the Hamiltonian

$$\boldsymbol{H}\boldsymbol{a}_n = E_n\boldsymbol{a}_n \quad \rightsquigarrow \quad \hat{H}|\Psi_n\rangle = E_n|\Psi_n\rangle \quad \text{with } |\Psi_n\rangle = |\boldsymbol{\Phi}\rangle\,\boldsymbol{a}_n\,. \tag{58}$$

Unfortunately, for an ab-initio Hamiltonian like (55) 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 $|\tilde{\boldsymbol{\Phi}}\rangle := \big(|\tilde{\Phi}_1\rangle,\dots,|\tilde{\Phi}_{\tilde{L}}\rangle\big)$, 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 eigenvectors

$$\tilde{\boldsymbol{H}}\tilde{\boldsymbol{a}}_n = \tilde{E}_n\tilde{\boldsymbol{a}}_n \quad \rightsquigarrow \quad |\tilde{\Psi}_n\rangle := |\tilde{\boldsymbol{\Phi}}\rangle\,\tilde{\boldsymbol{a}}_n \tag{59}$$

are related to those of H. The answer is surprisingly simple [6]: The eigenvalues of $\tilde{\boldsymbol{H}}$, ordered as $\tilde{E}_0 \le \tilde{E}_1 \le \dots \le \tilde{E}_{\tilde{L}-1}$, are variational with respect to those of H:

$$E_n \le \tilde{E}_n \quad \text{for } n \in \{0,\dots,\tilde{L}-1\}\,. \tag{60}$$

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

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 (60) with E_n being the eigenvalues of the L-dimensional Hamiltonian matrix $\boldsymbol{H}$. Being finite, we can use the same argument for $-\boldsymbol{H}$, obtaining

$$-E_{L-i} \leq -\tilde{E}_{\tilde{L}-i} \quad \text{for } i \in \{1,\ldots,\tilde{L}\}. \tag{61}$$

Taking the two inequalities together we obtain

$$E_n \leq \tilde{E}_n \leq E_{n+(L-\tilde{L})} \quad \text{for } n \in \{0,\ldots,\tilde{L}-1\}. \tag{62}$$

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

$$E_1 \leq \tilde{E}_1 \leq E_2 \leq \tilde{E}_2 \leq \cdots \leq \tilde{E}_L \leq E_{L+1}\,. \tag{63}$$

3.2 Matrix eigenvalue problem

For practical calculations we have to set up the Hamiltonian matrix $\tilde{\boldsymbol{H}} = \langle\tilde{\boldsymbol{\Phi}}|\hat{H}|\tilde{\boldsymbol{\Phi}}\rangle$ and the state vectors $\tilde{\boldsymbol{a}}$ for the chosen basis. This is particularly easy for a basis of Slater determinants constructed from a basis set of K orbitals $\{\varphi_k\,|k=0,\ldots,K-1\}$. The basis states are then the N-electron product states of $|\Phi_{k_1,\ldots,k_N}\rangle = c^\dagger_{k_N}\cdots c^\dagger_{k_1}|0\rangle$ with $k_1 < \cdots k_N$. We can write this more computer friendly as

$$|n_{K-1},\ldots,n_0\rangle = \prod_{k=0}^{K-1}\left(c^\dagger_k\right)^{n_k}|0\rangle \tag{64}$$

which is the occupation number representation with $n_k \in \{0,1\}$ and $\sum n_k = N$. It is natural to interpret the vector of occupation numbers as the binary representation of the integer $\sum_k 2^{n_k}$. This implies a natural ordering of the basis functions $|\Phi_l\rangle$. For the simple case of K=4 orbitals and N=2 electrons we obtain

i	(n_3,n_2,n_1,n_0)	state	l
0	0000		
1	0001		
2	0010		
3	0011	$c^\dagger_1 c^\dagger_0\|0\rangle = \|\Phi_1\rangle$	1
4	0100		
5	0101	$c^\dagger_2 c^\dagger_0\|0\rangle = \|\Phi_2\rangle$	2
6	0110	$c^\dagger_2 c^\dagger_1\|0\rangle = \|\Phi_3\rangle$	3
7	0111		
8	1000		
9	1001	$c^\dagger_3 c^\dagger_0\|0\rangle = \|\Phi_4\rangle$	4
10	1010	$c^\dagger_3 c^\dagger_1\|0\rangle = \|\Phi_5\rangle$	5
11	1011		
12	1100	$c^\dagger_3 c^\dagger_2\|0\rangle = \|\Phi_6\rangle$	6
13	1101		
14	1110		
15	1111		

The bit representation of the basis states also simplifies setting up the Hamiltonian matrix. Given the Hamiltonian in second quantization

$$\hat{H} = \sum_{n,m} T_{nm}\, c_n^\dagger c_m + \sum_{n'>n,\, m'>m} \underbrace{\big(U_{nn',mm'} - U_{n'n,mm'}\big)}_{=\breve{U}_{nn',mm'}} c_{n'}^\dagger c_n^\dagger c_m c_{m'} \tag{65}$$

the matrix element $\langle\Phi_l|\hat{H}|\Phi'_l\rangle$, with $|\Phi_{l'}\rangle = c^\dagger_{k'_N}\cdots c^\dagger_{l'_1}|0\rangle$, is given by

$$\sum_{n,m} T_{nm}\langle 0|c_{l_1}\cdots c_{l_N}\, c_n^\dagger c_m\, c^\dagger_{l'_N}\cdots c^\dagger_{l'_1}|0\rangle + \sum_{\substack{n'>n\\ m'>m}} \breve{U}_{nn',mm'}\langle 0|c_{l_1}\cdots c_{l_N}\, c_{n'}^\dagger c_n^\dagger c_m c_{m'}\, c^\dagger_{l'_N}\cdots c^\dagger_{l'_1}|0\rangle.$$

Anticommuting the operators coming from the operators, the matrix elements become overlaps of $N{+}1$ and $N{+}2$-electron product states, which, by (33) and (7), are just the determinants of the overlap matrices of the corresponding orbitals. When $\hat{H}$ is written in the same orbitals as the $|\Phi_l\rangle$, the overlap matrices simplify to permutation matrices with determinant ±1. In the occupation number representation, calculating this Fermi sign reduces to counting set bits. As an example we consider a simple hopping of an electron:

$$\begin{aligned} c_6^\dagger c_2|\Phi_{l(181)}\rangle &= c_6^\dagger c_2\, c_7^\dagger c_5^\dagger c_4^\dagger c_2^\dagger c_0^\dagger|0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger c_2 c_2^\dagger c_0^\dagger|0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger\big(1 - c_2^\dagger c_2\big) c_0^\dagger|0\rangle \\ &= (-1)^3 c_6^\dagger c_7^\dagger c_5^\dagger c_4^\dagger \cdot c_0^\dagger|0\rangle \\ = +|\Phi_{l(241)}\rangle &= (-1)^2 c_7^\dagger c_6^\dagger c_5^\dagger c_4^\dagger \cdot c_0^\dagger|0\rangle \end{aligned}$$

In the occupation number representation this becomes

$$\overset{\curvearrowright}{1\,0\,1\,1\,0\,1\,0\,1} \quad = \quad (-1)^c \;\; 1\,1\,1\,1\,0\,0\,0\,1 \tag{66}$$

where c is the count of set bits between the positions the electron hop. Note that a dedicated machine instruction, `popcnt`, for counting set bits is part of the x86 SSE4 instruction set, see also [7]

3.3 Dimension of the Hilbert space and sparseness

Setting up basis states and Hamiltonian matrix in this way, we can easily solve the many-body problem on our variational space by using any linear algebra library. This is the exact diagonalization approach. As discussed above, it gives us variational estimates of the ground and excited states. But there is a serious practical problem: the dimension of the many-body Hilbert space. For an N-electron problem with a basis set of K orbitals there are $K\,(K{-}1)\,(K{-}2)\cdots(K-(N{-}1))$ ways of picking N indices out of K. Since we only use one specific ordering of these indices, we have to divide by $N!$ to obtain the number of such determinants:

$$\dim \mathcal{H}_K^{(N)} = \frac{K!}{N!(K{-}N)!} = \binom{K}{N}. \tag{67}$$

Using Stirling's formula we see that for an N-electron problem this increases faster than exponentially with the size K of the basis set. This is the problem we face when converging the basis set for a finite system, e.g., a molecule. For solids we usually keep the number of orbitals per lattice site fixed, but scale to the thermodynamic limit, increasing the system size M while keeping the electron density N/M fixed. Also here the Hilbert space increases faster than exponentially. To give an impression of the problem we note that for $N = 25$ electrons and $K = 100$ orbitals the dimension already exceeds 10^{23}.

For exact diagonalization the problem gets even worse. Assuming we have a machine with 1 TeraBytes $= 2^{40}$ Bytes of RAM available. Using single precision (4 bytes) for the matrix elements, storing a matrix of dimension $(2^{40}/4)^{1/2} = 524\,288$ would already use up all memory. The dimension problem can be somewhat mitigated by exploiting symmetries: When the Hamiltonian commutes with the projection of the total spin, the number of up- and down-spin electrons is conserved separately. The N-electron Hamiltonian is then block diagonal in the sectors with fixed $N_\uparrow$ and $N_\downarrow$. The dimension of these blocks is significantly smaller than that of the full N-electron Hilbert space. Using the same orbital basis for each spin

$$\dim \mathcal{H}_{2K}^{(N_\uparrow,N_\downarrow)} = \binom{K}{N_\uparrow} \times \binom{K}{N_\downarrow}. \tag{68}$$

The S_z symmetry can be very easily implemented using the same ideas as introduced for the general case: just carry bit representations for the up- and down-spin electrons separately. In fact, when the total spin projection is conserved, we can distinguish electrons of different spin! Still, the Hilbert space of the single-band, half-filled Hubbard model with just 12 sites has dimension $853\,776$. Using further symmetries, if they exist, we could bring down the dimension somewhat further, however at the expense of considerable and problem-specific effort.

The key to going to larger systems is the realization that the vast majority of the elements of the Hamiltonian matrix is zero. This is quite easy to see. For the ab-initio Hamiltonian (55) with electron-electron repulsion, matrix elements between configurations that differ in more than two electron occupations vanish. Thus, for each configuration there may only be the diagonal element, $N \times (K-1)$ hopping terms, and $N(N-1)/2 \times (K-N)(K-N-1)/2$ pair-hopping terms. Thus the fraction of non-zero matrix elements of $\tilde{\boldsymbol{H}}$ to the total number is

$$\left(1 + N\left(1 + \frac{N-1}{2}\,\frac{(K-N-1)}{2}\right)(K-N)\right) \Big/ \binom{K}{N} \tag{69}$$

which, with increasing problem size, rapidly approaches zero. For the example of $N = 25$ electrons in $K = 100$ orbitals only 834 376 of the (over 10^{23}) matrix elements per row can be non-zero. This is the worst case. The sparsity of many-body Hamiltonians is even more pronounced when working in a tight-binding basis with short-ranged hopping and local electron-electron repulsion. Thus, many-body Hamiltonians are exceedingly sparse and the more so the larger the problem. They are therefore ideally suited for approaches like the Lanczos method, that are based on matrix-vector products, which for the sparse matrices scale close to linearly in the matrix dimension.

3.4 Non-interacting electrons

Even when considering a system of N non-interacting electrons we have to solve the large matrix eigenvalue problem (57). Writing the non-interacting Hamiltonian in the basis used for the CI expansion (56) 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 basis for the CI expansion. If we choose the eigenstates of the single-electron matrix H_{nm} as basis, the second-quantized Hamiltonian is

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

In this basis all off-diagonal matrix elements vanish and the CI Hamiltonian (57) is diagonal. Thus all $\binom{K}{N}$ eigenstates are Slater determinants

$$|\Phi_{\boldsymbol{n}}\rangle = c_{n_N}^\dagger \cdots c_{n_1}^\dagger |0\rangle \qquad \text{with eigenenergy} \quad E_{\boldsymbol{n}} = \sum_i \varepsilon_{n_i} \,. \tag{70}$$

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 (57) and at the same time look for the basis set (of given size) that minimizes the variational energy. The simplest example, where the many-body basis consists of a single Slater determinant, is the Hartree-Fock method.

4 Jordan-Wigner representation

The breakthrough in the development of quantum mechanics was Heisenberg's insight that position and momentum of a particle must be represented by non-commuting matrices [8]. The canonical commutation relations $[\hat{r}_i,\, \hat{p}_j] = \mathrm{i}\hbar\delta_{ij}$ are the essence of quantization. In order to develop a quantum theory of radiation, Dirac quantized a plane wave $\vec{E}_{\vec{k}}(\vec{r},t) = \vec{E}_{\vec{k}}\, \mathrm{e}^{\mathrm{i}(\vec{k}\cdot\vec{r}-\omega t)}$ using the same idea, describing the wave by its energy E and phase $\theta = -\omega t$, postulating matrices with canonical commutation relation $[E,t] = \mathrm{i}\hbar$. In analogy to Heisenberg's first quantization, this step is called *second* quantization. It is quite plausible, considering the time-dependent Schrödinger equation: $\mathrm{i}\hbar\frac{d}{dt} = E$. It follows from the commutation relation $[E,\, \mathrm{e}^{-\mathrm{i}\omega t}] = \hbar\omega\, \mathrm{e}^{-\mathrm{i}\omega t}$ that the operator $\mathrm{e}^{\mathrm{i}\theta}$ adds a photon to the radiation field: $E\mathrm{e}^{\mathrm{i}\theta} = \mathrm{e}^{\mathrm{i}\theta}(E{+}\hbar\omega)$.
Instead of working with these operators directly, Dirac relates them to the expansion coefficient b_k of a many-photon wave function in the basis of plane waves and requires that $|b_k|^2 = b_k^\dagger b_k$ gives the occupation $N_k{=}E/\hbar\omega_k$ of mode k. This means that, in second quantization, the expansion coefficients become operators, e.g.,

$$b_k = \mathrm{e}^{-\mathrm{i}\theta_k} N_k^{1/2} \quad \text{and} \quad b_k^\dagger = N_k^{1/2} \mathrm{e}^{\mathrm{i}\theta_k} \tag{71}$$

with $b_k^\dagger b_k = N_k$ and commutation relation

$$\left[b_k,\, b_k^\dagger\right] = \mathrm{e}^{-\mathrm{i}\theta_k} N_k \mathrm{e}^{\mathrm{i}\theta_k} - N_k = \mathrm{e}^{-\mathrm{i}\theta_k}\mathrm{e}^{\mathrm{i}\theta_k}(N_k{+}1) - N_k = 1\,. \tag{72}$$

Pascual Jordan realized [10] that the central part of Dirac's argument is the commutation relation $[N_k,\, \mathrm{e}^{\mathrm{i}\theta_k}] = \mathrm{e}^{\mathrm{i}\theta_k}$, showing that $\mathrm{e}^{\mathrm{i}\theta}$ acts as a ladder operator. To derive a similar description for electrons, he implemented the Pauli principle by restricting occupations to the values 0 and 1. In its eigenbasis, N_k can then be written as a $2{\times}2$ matrix

$$N = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_z{+}1) \tag{73}$$

so that the ladder operators, fulfilling $\left[N,\, \mathrm{e}^{\pm\mathrm{i}\theta}\right] = \pm\mathrm{e}^{\pm\mathrm{i}\theta}$, take the form

$$\mathrm{e}^{\mathrm{i}\theta} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_x{+}\mathrm{i}\sigma_y) \quad \text{and} \quad \mathrm{e}^{-\mathrm{i}\theta} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_x{-}\mathrm{i}\sigma_y) \tag{74}$$

and, using $N^2 = N \rightsquigarrow N^{1/2} = N$, the creators and annihilators (71) become

$$d^\dagger = N^{1/2}\,\mathrm{e}^{\mathrm{i}\theta} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_x{+}\mathrm{i}\sigma_y) \tag{75}$$

$$d = \mathrm{e}^{-\mathrm{i}\theta} N^{1/2} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \frac{1}{2}(\sigma_x{-}\mathrm{i}\sigma_y). \tag{76}$$

The resulting Paulions do not fulfill a canonical *commutation* relation ($[d^\dagger,\, d] = \sigma_z$), instead their *anti*-commutator is analogous to (72)

$$d\,d^\dagger + d^\dagger d = \left\{d,\, d^\dagger\right\} = 1\,. \tag{77}$$

Using $S^\pm = S^x \pm \mathrm{i}S^y = (\sigma_x \pm \mathrm{i}\sigma_y)/2$ and $S^z = \sigma_z/2$, we see from the simple mapping

paulion	$\vert\,0\,\rangle$	$\vert\,1\,\rangle$	$d^\dagger$	d	N
spin-½	$\vert\downarrow\rangle$	$\vert\uparrow\rangle$	S^+	S^-	$S^z{+}\frac{1}{2}$

(78)

that a paulions behaves just like a spin-½. Extending the analogy to more modes/spins, we obtain the algebra of hard-core bosons [12], where operators involving different spins commute: $[d_i,\, d_j] = 0 = [d_i^\dagger,\, d_j^\dagger]$, and $[d_i,\, d_j^\dagger] = 0$ for $i{\neq}j$, while on the same site they fulfill the canonical anticommutation relations

$$\left\{d_i,\, d_i^\dagger\right\} = 1 \quad \text{and} \quad \left\{d_i,\, d_i\right\} = 0 = \left\{d_i^\dagger,\, d_i^\dagger\right\}. \tag{79}$$

While hard-core bosons always fulfill the Pauli principle: $\left(d_i^\dagger\right)^2 = 0$, they change character under basis transformations. This is most easily seen in the case of two sites: Transforming from d_1 and d_2 to $d_\pm = (d_1 \pm d_2)/\sqrt{2}$, we see that in the new basis the operators on different modes need no longer commute

$$\left[d_+,\, d_-^\dagger\right] = \frac{1}{2}\Big(\left[d_1, d_1^\dagger\right] - \left[d_2,\, d_2^\dagger\right]\Big) = N_2 - N_1\,,$$

where we have used $[A,\,B] = \{A,\,B\} - 2BA$. The problem arises from the on-site anticommutators mixing into the off-site commutators.
This can be fixed by requiring fermionic anticommutation relations for all operators

$$\left\{c_i,\,c_j^\dagger\right\} = \delta_{ij} \quad \text{and} \quad \left\{c_i,\,c_j\right\} = 0 = \left\{c_i^\dagger,\,c_j^\dagger\right\} \tag{80}$$

which, just like the bosonic commutators, remain unchanged under (unitary) basis transformations $c_{\beta_i}^\dagger = \sum_\mu c_{\alpha_\mu}^\dagger U_{\mu i}$:

$$[c_{\beta_i}, c_{\beta_j}^\dagger]_\pm = \sum_{\mu,\nu} \overline{U_{\mu i}}\, U_{\nu j} \underbrace{[c_{\alpha_\mu}, c_{\alpha_\nu}^\dagger]_\pm}_{=\delta_{\mu,\nu}} = \sum_\mu (U^\dagger)_{i\mu} U_{\mu j} = 1\,. \tag{81}$$

This was realized by Jordan and Wigner [11]. To make operators on different sites anticommute, they introduced operators to keep track of the Fermi sign encountered, e.g., when calculating hopping terms in the occupation number representation as in (66). The Fermi string operator $\prod_{k=1}^{i-1}(-\sigma_z^k)$ collects a factor -1 for every occupied mode (i.e., spin up) below the mode the ladder operator acts on

$$c_i^\dagger = S_i^+ \otimes (-2S_{i-1}^z) \otimes \cdots \otimes (-2S_1^z) \quad \text{and} \quad c_i = S_i^- \otimes (-2S_{i-1}^z) \otimes \cdots \otimes (-2S_1^z) \tag{82}$$

Since $\left(\sigma_z\right)^2 = 1$ this does not change the anticommutation relations between operators on the same sites, e.g.,

$$\left\{c_i,\,c_i^\dagger\right\} = \left\{S_i^-,\,S_i^+\right\} \otimes \left(-\sigma_z^{i-1}\right)^2 \otimes \cdots \otimes \left(-\sigma_z^1\right)^2 = 1 \tag{83}$$

while the operators on different sites now anticommute as desired, e.g., for $i{>}j$

$$\left\{c_i,\,c_j^\dagger\right\} = S_i^- \otimes \left(-\sigma_z^{i-1}\right) \otimes \cdots \otimes \left(-\sigma_z^{j+1}\right) \otimes \underbrace{\left\{S_j^+, -\sigma_z^i\right\}}_{=0} \otimes \left(-\sigma_z^{j-1}\right)^2 \otimes \cdots \left(-\sigma_z^1\right)^2 = 0\,. \tag{84}$$

Thus the operators c_i and $c_j^\dagger$ fulfill canonical anticommutation relations, the form of which remains unchanged under basis transformations, cf. (81). It follows from (83) that also for paulions the hard-core constraint $N_i^2 = N_i$ is invariant under basis transformations.
We note in passing that the Pauli principle has interesting consequences for the classical description of electrons first noticed by Rudolf Peierls [6]: Given the particle-wave duality in quantum mechanics one may wonder why, in the classical limit, electrons behave not like radiation but as particles. This may be seen as a consequence of the uncertainty relation between particle number and phase following from $[N, \theta] = -\mathrm{i}$, which formed the basis of second quantization. A classical wave has a well defined phase, which can only be obtained when the uncertainty in N is large. This condition can be fulfilled for bosons in the limit of large average number of photons. For fermions, however, the uncertainty in N is severely limited by the Pauli principle, making it impossible to get a classical wave with well defined phase.

Such arguments based on a phase operator, as well as the derivation of second quantization sketched above have, however, to be taken with caution, as working with phase and angle operators is not straightforward [13]. We can see that there is something not quite right with the phase operator θ already from (74), which shows that $\mathrm{e}^{\mathrm{i}\theta}$ is *not* unitary, i.e., that θ is not hermitian. The same is true for the corresponding ladder operators in the bosonic case. The reason is that the number of particles is bounded from below, which means that there is a lowest state that is annihilated by $\mathrm{e}^{-\mathrm{i}\theta}$, so that this operator cannot have an inverse.

This problem no longer appears in relativistic quantum mechanics: the second-quantization commutation relation $[E,\, t] = \mathrm{i}\hbar$ is related to the first-quantization $[r_i,\, p_i] = \mathrm{i}\hbar\delta_{ij}$ by Lorentz invariance. Correspondingly, the relativistic energy spectrum is not bounded from below, but has hole states of negative energy.

Non-relativistic second quantization, on the other hand, does not need these commutation relations, since it is nothing but an extremely economical way of representing states. The bosonic commutation relations are just those for the ladder operators $a = \big(x/x_0 + \mathrm{i}x_0 p_x/\hbar\big)/\sqrt{2}$ of a harmonic oscillator with $x_0 = \sqrt{\hbar/m\omega}$, for which they are, in fact, just a consequence of the position-momentum commutator

$$\mathrm{i}\hbar = [x,\, p_x] = \Big[\frac{x_0}{\sqrt{2}}\big(a{+}a^\dagger\big),\, -\frac{\mathrm{i}\hbar}{\sqrt{2}x_0}\big(a{-}a^\dagger\big)\Big] = -\frac{\mathrm{i}\hbar}{2}\big[a{+}a^\dagger,\, a{-}a^\dagger\big] = +\mathrm{i}\hbar\big[a,\, a^\dagger\big].$$

For fermions there is a simple relation between an N-electron wave function $\Psi(x_1, \ldots, x_N)$ and the corresponding N-electron state $|\Psi\rangle$:

$$|\Psi\rangle = \frac{1}{\sqrt{N!}}\int dx_1 \cdots dx_N\, \Psi(x_1, \ldots, x_N)\, \hat\Psi^\dagger(x_N) \cdots \hat\Psi^\dagger(x_1)|0\rangle \tag{85}$$

$$\Psi(x_1, \ldots, x_N) = \frac{1}{\sqrt{N!}}\langle 0|\hat\Psi(x_1) \cdots \hat\Psi(x_N)|\Psi\rangle \tag{86}$$

with unitarity relations

$$\frac{1}{N!}\int dx_1 \cdots dx_N\, \hat\Psi^\dagger(x_N) \cdots \hat\Psi^\dagger(x_1)|0\rangle\langle 0|\hat\Psi(x_1) \cdots \hat\Psi(x_N) = \mathbb{1}_N \tag{87}$$

and

$$\langle 0|\hat\Psi(x_1) \cdots \hat\Psi(x_N)\, \hat\Psi^\dagger(x'_N) \cdots \hat\Psi^\dagger(x'_1)|0\rangle = \begin{vmatrix} \delta(x_1{-}x'_1) & \cdots & \delta(x_1{-}x'_N) \\ \vdots & & \vdots \\ \delta(x_N{-}x'_1) & \cdots & \delta(x_N{-}x'_N) \end{vmatrix} \tag{88}$$

so that normalizations are kept

$$\int dx_1 \cdots dx_N\, \big|\Psi(x_1, \ldots, x_N)\big|^2 = 1 \quad \Leftrightarrow \quad \langle\Psi|\Psi\rangle = 1\,. \tag{89}$$

Note that first-quantized wave functions involve factors $1/\sqrt{N!}$, since the normalization integral is over all permutations of the coordinates, while integration over a single domain $x_1 < \cdots < x_n$ (assuming a suitable ordering of the coordinates) already contains all relevant information.

4.1 Spins, hard-core bosons, and fermions

The equivalence between spin operators and hard-core bosons makes it possible to rewrite spin-½ Heisenberg models like

$$J\sum_{\langle ij\rangle}\vec{S}_i\cdot\vec{S}_j \tag{90}$$

using $S^x=(S^+ + S^-)/2$, $S^y=(S^+ - S^-)/2\mathrm{i}$, and (78) as dictionary, into

$$J\sum_{\langle ij\rangle}\frac{1}{2}\Big(d_i^\dagger d_j + d_i d_j^\dagger\Big) + \Big(d_i^\dagger d_i - \tfrac{1}{2}\Big)\Big(d_j^\dagger d_j - \tfrac{1}{2}\Big), \tag{91}$$

where an (anti)ferromagnetic coupling of the spins corresponds to a (repulsive) attractive interaction between the hard-core bosons, while the transverse components of the spin interaction corresponds to a hopping of paulions. Thus, e.g., the XY-model

$$\sum_{\langle ij\rangle} J_x\, S_i^x S_j^x + J_y\, S_i^y S_j^y \tag{92}$$

maps to the non-interacting hard-core boson model

$$\sum_{\langle ij\rangle}\frac{J_x+J_y}{4}\Big(d_i^\dagger d_j + d_i d_j^\dagger\Big) + \frac{J_x-J_y}{4}\Big(d_i^\dagger d_j^\dagger + d_i d_j\Big). \tag{93}$$

For the XX-model, $J_x{=}J_y$, there are just hopping terms, for the general XY-model there are also pairing terms. The opposite limit, $J_x{=}0{=}J_y$, the Ising model, here with transverse field

$$J\sum_{\langle ij\rangle} S_i^z S_j^z + H\sum_i S_i^x \tag{94}$$

seems at first to necessarily lead to a model of interacting hard-core bosons. We can, however, easily rotate the spin-direction so that the external field points in z direction, to obtain a hard-core boson model with only one-body terms

$$\frac{J}{4}\sum_{\langle ij\rangle}\big(d_i^\dagger + d_i\big)\big(d_j^\dagger + d_j\big) + H\sum_i\Big(d_i^\dagger d_i - \tfrac{1}{2}\Big)\,. \tag{95}$$

where the first term is just (93) with $J_y=0$.

It is tempting to try to solve these non-interacting hard-core boson Hamiltonians by diagonalizing the one-body matrix and using a Bogoliubov transformation. As we have seen, however, the commutation relations between paulions are not invariant under basis changes, so that the eigenstates are not simple product states as in (70). This only works for fermions, i.e., when we multiply the Pauli operators with Fermi strings to obtain fermions $c_i^\dagger = d_i^\dagger\prod_{k=1}^{i-1}(-1)^{n_k}$ with $n_k = d_k^\dagger d_k$ or, equivalently, $d_i^\dagger = c_i^\dagger\prod_{k=1}^{i-1}(-1)^{n_k}$ with $n_k = c_k^\dagger c_k$. Since the sign operators act on different modes than the creator/annihilator, they commute with the Fermi strings. Rewriting the paulion models we thus get terms like

$$d_i^\dagger d_j = c_i^\dagger\prod_{k=1}^{i-1}(-1)^{n_k}\prod_{k=1}^{j-1}(-1)^{n_k}c_j = c_i^\dagger\prod_{k=\min(i,j)}^{\max(i,j)-1}(-1)^{n_k}\,c_j\,. \tag{96}$$

so that the simple hopping or pairing terms are complicated by the product of spin operators. Things remain simple, however, for terms involving the same site

$$d_i^\dagger d_i = c_i^\dagger \sum_{k=i}^{i-1} (-1)^{n_k} c_i = c_i^\dagger c_i \tag{97}$$

or adjacent sites (just consider the cases n_i=0 or 1 separately)

$$d_{i+1}^\dagger d_i = c_{i+1}^\dagger (-1)^{n_i} c_i = c_{i+1}^\dagger c_i \quad \text{and} \quad d_{i+1}^\dagger d_i^\dagger = c_{i+1}^\dagger (-1)^{n_i} c_i^\dagger = -c_{i+1}^\dagger c_i^\dagger \tag{98}$$

and, taking the adjoint

$$d_i^\dagger d_{i+1} = \left(d_{i+1}^\dagger d_i\right)^\dagger = c_i^\dagger c_{i+1} \quad \text{and} \quad d_i d_{i+1} = \left(d_{i+1}^\dagger d_i^\dagger\right)^\dagger = -c_i c_{i+1}\,. \tag{99}$$

For a one-dimensional spin chain, the XY-model or the Ising model with transverse field can thus be written as a model of non-interacting fermions, e.g., (93) becomes

$$\sum_j \frac{J_x+J_y}{4}\left(c_{j+1}^\dagger c_j + c_j^\dagger c_{j+1}\right) + \frac{J_x-J_y}{4}\left(-c_{j+1}^\dagger c_j^\dagger - c_j c_{j+1}\right) \tag{100}$$

and, since now we are dealing with non-interacting fermions, solved in the usual way [14].

A Appendix

A.1 Atomic units

Practical electronic structure calculations are usually done in atomic units, a.u. for short. While the idea behind the atomic units is remarkably simple, in practice there is often some confusion when trying to convert to SI units. We therefore give a brief explanation.
The motivation for introducing atomic units is to simplify the equations. For example, in SI units the Hamiltonian of a hydrogen atom is

$$H = -\frac{\hbar^2}{2m_e}\nabla^2 - \frac{e^2}{4\pi\epsilon_0\, r}\ . \tag{101}$$

When we implement such an equation in a computer program, we need to enter the numerical values of all the fundamental constants. We can avoid this by using a system of units in which the *numerical values* of the electron mass m_e, the elementary charge e, the Planck-constant $\hbar$, and the dielectric constant $4\pi\epsilon_0$ are all equal to one. In these units the above equation can be programmed as

$$H = -\frac{1}{2}\nabla^2 - \frac{1}{r}\ . \tag{102}$$

This immediately tells us: 1 a.u. mass $= m_e$ and 1 a.u. charge $= e$. To complete the set of basis units we still need the atomic unit of length, which we call a_0, and of time, t_0. To find the values of a_0 and t_0 we write $\hbar$ and $4\pi\epsilon_0$ (using simple dimensional analysis) in atomic units: $\hbar = 1\, m_e a_0^2/t_0$ and $4\pi\epsilon_0 = 1\, t_0^2 e^2/(m_e a_0^3)$. Solving this system of equations, we find

1 a.u. length	$= a_0 =$	$4\pi\epsilon_0\hbar^2/m_e e^2$	$\approx 5.2918\cdot 10^{-11}$	m
1 a.u. mass	$= m_e =$		$\approx 9.1095\cdot 10^{-31}$	kg
1 a.u. time	$= t_0 =$	$(4\pi\epsilon_0)^2\hbar^3/m_e e^4$	$\approx 2.4189\cdot 10^{-17}$	s
1 a.u. charge	$= e =$		$\approx 1.6022\cdot 10^{-19}$	C

The atomic unit of length, a_0, is the Bohr radius. As the dimension of energy is mass times length squared divided by time squared, its atomic unit is $m_e a_0^2/t_0^2 = m_e e^4/(4\pi\epsilon_0)^2\hbar^2$. Because of its importance the atomic unit of energy has a name, the Hartree. One Hartree is minus twice the ground-state energy of the hydrogen atom, $1\,\mathrm{Ha} \approx 27.211\,\mathrm{eV}$.
It would be tempting to try to set the numerical value of all fundamental constants to unity. But this must obviously fail, as the system of equations to solve becomes overdetermined when we try to prescribe the numerical values of constants that are not linearly independent in the space of basis units. Thus, we cannot, e.g., choose also the speed of light to have value one, as would be practical for relativistic calculations. Instead, in atomic units it is given by $c\, t_0/a_0 = 4\pi\varepsilon_0\hbar c/e^2 = 1/\alpha$, where α is the fine structure constant. Thus $c = \alpha^{-1}$ a.u. ≈ 137 a.u. The Bohr magneton is $\mu_B = 1/2$ a.u. The Boltzmann constant k_B, on the other hand, is independent of the previous constants. Setting its value to one fixes the unit of temperature to 1 a.u. temperature $= m_e e^4/(4\pi\epsilon_0)^2\hbar^2 k_B = \mathrm{Ha}/k_B \approx 3.158\cdot 10^5\,\mathrm{K}$.

A.2 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

$$\mathbb{1} = \textstyle\sum_{k,l} |\chi_k\rangle (S^{-1})_{kl} \langle\chi_l| \,. \tag{103}$$

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

$$\sum_k \underbrace{\langle\chi_i|H|\chi_k\rangle}_{=:H_{ik}} \underbrace{\sum_l (S^{-1})_{kl}\langle\chi_l|v\rangle}_{=:v_k} \langle\chi_i|\hat{H}|v\rangle = \varepsilon\langle\chi_i|v\rangle = \varepsilon \sum_k \underbrace{\langle\chi_i|\chi_k\rangle}_{=S_{ik}} \underbrace{\sum_l (S^{-1})_{kl}\langle\chi_l|v\rangle}_{=v_k} \,. \tag{104}$$

Collecting all components, this becomes the generalized eigenvalue problem $\boldsymbol{H}\boldsymbol{v} = \varepsilon \boldsymbol{S}\boldsymbol{v}$. From the solution $\boldsymbol{v}$ 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 $\boldsymbol{T}$ such that

$$|\varphi_n\rangle := \textstyle\sum_m |\chi_m\rangle \, T_{mn} \qquad \text{with} \quad \langle\varphi_n|\varphi_m\rangle = \delta_{mn} \,. \tag{105}$$

This implies that $\boldsymbol{T}^\dagger \boldsymbol{S}\boldsymbol{T} = \mathbb{1}$, or equivalently $\boldsymbol{S}^{-1} = \boldsymbol{T}\boldsymbol{T}^\dagger$. This condition does not uniquely determine $\boldsymbol{T}$. 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 $\boldsymbol{T}$ that minimizes

$$\begin{aligned} \sum_n \big\| |\varphi_n\rangle - |\chi_n\rangle \big\|^2 &= \sum_n \Big\| \sum_m |\chi_m\rangle (T_{mn} - \delta_{mn}) \Big\|^2 \\ &= \mathrm{Tr}\,(\boldsymbol{T}^\dagger - \mathbb{1})\, \boldsymbol{S}\, (\boldsymbol{T} - \mathbb{1}) = \mathrm{Tr}\,(\underbrace{\boldsymbol{T}^\dagger \boldsymbol{S}\boldsymbol{T}}_{=\mathbb{1}} - \boldsymbol{T}^\dagger\boldsymbol{S} - \boldsymbol{S}\boldsymbol{T} + \boldsymbol{S}) \,. \end{aligned} \tag{106}$$

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

$$0 \stackrel{!}{=} \mathrm{Tr}\,(+i\boldsymbol{M}\boldsymbol{T}^\dagger\boldsymbol{S} - i\boldsymbol{S}\boldsymbol{T}\boldsymbol{M}) = i\,\mathrm{Tr}\,(\boldsymbol{T}^\dagger\boldsymbol{S} - \boldsymbol{S}\boldsymbol{T})\boldsymbol{M}, \tag{107}$$

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

$$\frac{1}{2}\,\mathrm{Tr}\,(\boldsymbol{M}^2\boldsymbol{S}^{1/2} + \boldsymbol{S}^{1/2}\boldsymbol{M}^2) > 0 \tag{108}$$

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

$$\boldsymbol{T}_{\text{Löwdin}} = \boldsymbol{S}^{-1/2} \tag{109}$$

minimizes the modification of the basis vectors due to orthogonalization.

A.3 Pauli matrices

The spin matrices were defined by Pauli [16] as

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

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

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

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

$$\big(\vec{a}\cdot\vec{\sigma}\big)\big(\vec{b}\cdot\vec{\sigma}\big) = \sum a_j b_k\, \sigma_j\sigma_k = \big(\vec{a}\cdot\vec{b}\big)\mathbb{1} + \mathrm{i}\big(\vec{a}\times\vec{b}\big)\cdot\vec{\sigma}\,. \tag{110}$$

From the products follow the familiar commutation relations $[\sigma_x,\, \sigma_y] = 2\mathrm{i}\sigma_z$ (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

$$\begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} = M = m_0\mathbb{1} + \vec{m}\cdot\vec{\sigma} = \begin{pmatrix} m_0{+}\,m_z & m_x{-}\mathrm{i}m_y \\ m_x{+}\mathrm{i}m_y & m_0{-}\,m_z \end{pmatrix} \tag{111}$$

with $2m_0 = m_{11}{+}m_{22}$, $2m_z = m_{11}{-}m_{22}$, $2m_x = m_{12}{+}m_{21}$, and $2m_x = \mathrm{i}(m_{12}{-}m_{21})$, which can be written as $2m_j = \operatorname{Tr} M\sigma_j$, with $\sigma_0 := \mathbb{1}$. When the m_0 and $\vec{m}$ are real, M is hermitian. Matrix products are easily evaluated using (110). 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

$$M = m_0\mathbb{1} + \vec{m}\cdot\vec{\sigma} \quad \rightsquigarrow \quad M^{-1}{=}(m_0\mathbb{1} - \vec{m}\cdot\vec{\sigma})/\det M. \tag{112}$$

For a unitary matrix $U{=}\,u_0{+}\vec{u}{\cdot}\vec{\sigma}$ with $\det U = 1$ we then see from $U^\dagger = u_0^*{+}\vec{u}^{\,*}{\cdot}\vec{\sigma} \overset{!}{=} u_0{-}\vec{u}{\cdot}\vec{\sigma}$ that u_0 must be real and $\vec{u} = \mathrm{i}\vec{n}$ 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 $U{\in}\,\mathrm{SU}(2)$ can be written, using $(\hat{n}\cdot\vec{\sigma})^2{=}\,\mathbb{1}$ from (110) in the power series,

$$U_{\hat{n},\alpha} = \cos\alpha\,\mathbb{1} + \mathrm{i}\sin\alpha\ (\hat{n}\cdot\vec{\sigma}) = \exp\big(\mathrm{i}\alpha\,\hat{n}\cdot\vec{\sigma}\big). \tag{113}$$

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

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 2×2 matrix with zero trace. By the cyclic property of the trace $\mathrm{Tr}\, U(\vec{a}\cdot\vec{\sigma})U^\dagger = \mathrm{Tr}\,\vec{a}\cdot\vec{\sigma}$, 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, $U\big((c\vec{a}+\vec{b})\cdot\vec{\sigma}\big)U^\dagger = cU(\vec{a}\cdot\vec{\sigma})U^\dagger + U(\vec{b}\cdot\vec{\sigma})U^\dagger$, and preserves the inner product

$$\vec{a}\cdot\vec{b} = \tfrac{1}{2}\,\mathrm{Tr}(\vec{a}\cdot\vec{\sigma})(\vec{b}\cdot\vec{\sigma}) = \tfrac{1}{2}\,\mathrm{Tr}\,U(\vec{a}\cdot\vec{\sigma})U^\dagger\,U(\vec{b}\cdot\vec{\sigma})U^\dagger = \tfrac{1}{2}\,\mathrm{Tr}(\vec{a}_U\cdot\vec{\sigma})(\vec{b}_U\cdot\vec{\sigma}) = \vec{a}_U\cdot\vec{b}_U$$

so that it must be a proper rotation, $\vec{a}_U = R_U\,\vec{a}$ with $R_U\in\mathrm{SO}(3)$. To identify which rotation, we consider the special case $\vec{a}_\parallel = c_a\hat{n}$ for which, by (110), $\vec{a}_\parallel\cdot\vec{\sigma}$ 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 ϑ, we consider a unit vector $\hat{a}_\perp$ perpendicular to $\hat{n}$, for which, using $(\hat{a}_\perp\cdot\vec{\sigma})(\hat{n}\cdot\vec{\sigma}) = \mathrm{i}(\hat{a}_\perp\times\hat{n})\cdot\vec{\sigma}$ and $\mathrm{Tr}\,\vec{v}\cdot\vec{\sigma} = 0$,

$$\begin{aligned}
\cos\vartheta = \hat{a}_\perp\cdot R_U\hat{a}_\perp &= \tfrac{1}{2}\,\mathrm{Tr}(\hat{a}_\perp\cdot\vec{\sigma})U(\hat{a}_\perp\cdot\vec{\sigma})U^\dagger \\
&= \tfrac{1}{2}\,\mathrm{Tr}(\hat{a}_\perp\cdot\vec{\sigma})(\cos\alpha + \mathrm{i}\sin\alpha\,(\hat{n}\cdot\vec{\sigma}))\;\;(\hat{a}_\perp\cdot\vec{\sigma})(\cos\alpha - \mathrm{i}\sin\alpha\,(\hat{n}\cdot\vec{\sigma})) \\
&= \tfrac{1}{2}\,\mathrm{Tr}\big(\cos\alpha\,\hat{a}_\perp - \sin\alpha\,(\hat{a}_\perp\times\hat{n})\big)\cdot\vec{\sigma}\quad\big(\cos\alpha\,\hat{a}_\perp + \sin\alpha\,(\hat{a}_\perp\times\hat{n})\big)\cdot\vec{\sigma} \\
&= \big(\cos\alpha\,\hat{a}_\perp - \sin\alpha\,(\hat{a}_\perp\times\hat{n})\big)\cdot\big(\cos\alpha\,\hat{a}_\perp + \sin\alpha\,(\hat{a}_\perp\times\hat{n})\big) = (\cos\alpha)^2 - (\sin\alpha)^2 = \cos 2\alpha
\end{aligned}$$

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 $\mathrm{SU}(2)$ 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 $\mathrm{SU}(2)$ induce the same rotation $R_{-U} = R_U\in\mathrm{SO}(3)$.

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

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

from which we easily read off the eigenvalues

$$\varepsilon_\pm = m_0 \pm \|\vec{m}\| = \frac{m_{11}{+}m_{22}}{2} \pm \sqrt{\frac{(m_{11}{-}m_{22})^2}{4} + |m_{12}|^2} = \tfrac{1}{2}\,\mathrm{Tr}\,M \pm \sqrt{\big(\tfrac{1}{2}\,\mathrm{Tr}\,M\big)^2 - \det M},$$

while the eigenvalues are the columns vectors of $U^\dagger = \big(\boldsymbol{v}_+,\,\boldsymbol{v}_-\big)$

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

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

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

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

$$\boldsymbol{v}_+ = \begin{pmatrix}\cos\frac{\vartheta}{2}\\ +\mathrm{e}^{+\mathrm{i}\varphi}\sin\frac{\vartheta}{2}\end{pmatrix}\;\text{and}\;\boldsymbol{v}_- = \begin{pmatrix}-\mathrm{e}^{-\mathrm{i}\varphi}\sin\frac{\vartheta}{2}\\ \cos\frac{\vartheta}{2}\end{pmatrix}\;\text{with}\;\begin{array}{l}\varphi = \arg(m_{21}) = -\arg(m_{12})\\ \vartheta = \arccos\dfrac{m_{11}{-}m_{22}}{\varepsilon_+{-}\varepsilon_-}\end{array}. \tag{114}$$

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

A.4 Some useful commutation relations

Commuting an operator through a product of operators is straightforward

$$\begin{aligned} AB_1\cdots B_N &= [A,\,B_1]\,B_2\cdots B_N + B_1AB_2\cdots B_N \\ &= [A,\,B_1]\,B_2\cdots B_N + B_1[A,\,B_2]\cdots B_N + B_1B_2A\cdots B_N \\ &\;\;\vdots \\ &= \sum_{n=1}^{N}\prod_{i=1}^{n-1} B_i\;\;[A,\,B_n]\prod_{i=n+1}^{N} B_i + B_1\cdots B_NA \end{aligned}$$

while, working analogously, anticommuting introduces alternating signs

$$\begin{aligned} AB_1\cdots B_N &= \{A,\,B_1\}\,B_2\cdots B_N - B_1AB_2\cdots B_N \\ &= \{A,\,B_1\}\,B_2\cdots B_N - B_1\{A,\,B_2\}\cdots B_N - B_1B_2A\cdots B_N \\ &\;\;\vdots \\ &= \sum_{n=1}^{N}(-1)^{n-1}\prod_{i=1}^{n-1} B_i\;\;\{A,\,B_n\}\prod_{i=n+1}^{N} B_i + (-1)^N B_1\cdots B_NA \end{aligned}$$

The following special cases are particularly useful

$$\begin{aligned} [A,\,BC] &= [A,\,B]\,C \,+\, B\,[A,\,C] \\ &= \{A,\,B\}C - B\{A,\,C\} \end{aligned}$$

$$\begin{aligned} [AB,\,C] &= A\,[B,\,C] \,+\, [A,\,C]\,B \\ &= A\{B,\,C\} - \{A,\,C\}B \end{aligned}$$

$$\begin{aligned} [AB,\,CD] &= A\,[B,\,C]\,D + AC\,[B,\,D] \,+\, [A,C]\,DB + C\,[A,\,D]\,B \\ &= A\{B,\,C\}D - AC\{B,\,D\} + \{A,C\}DB - C\{A,\,D\}B \end{aligned}$$

Important examples are $\left[c_i^\dagger c_j,\,c_\gamma^\dagger\right] = \langle j|\gamma\rangle\,c_i^\dagger$ and $\left[c_i^\dagger c_j,\,c_\gamma\right] = -\langle i\,|\gamma\rangle\,c_j$.
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] E. Koch: *Many-Electron States*, in [17]

[6] R. Peierls: *Surprises in Theoretical Physics* (Princeton University Press, 1979)

[7] H.S. Warren: *Hacker's Delight*, 2nd ed. (Addison-Wesley, 2013)

[8] W. Heisenberg: *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen,* Z. Phys. **33**, 879 (1925)

[9] P.M.A. Dirac: *The Quantum Theory of the Emission and Absorption of Radiation,*
Proc. Roy. Soc. Lond. A **114**, 243 (1927)

[10] P. Jordan: *Zur Quantenmechanik der Gasentartung,* Z. Phys. **44**, 473 (1927)

[11] P. Jordan and E. Wigner: *Über das Paulische Äquivalenzverbot,* Z. Phys. **47**, 631 (1928)

[12] T. Matsubara and H. Matsuda: *A Lattice Model of Liquid Helium, I,*
Prog. Theor. Phys. **16**, 569 (1956)

[13] P. Carruthers and M.M. Nieto: *Phase and Angle Variables in Quantum Mechanics,*
Rev. Mod. Phys. **40**, 411 (1968)

[14] E. Lieb, T. Schulz, and D. Mattis: *Two Soluble Models of an Antiferromagnetic Chain,*
Ann. Phys. **16**, 407 (1961)

[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,
Modeling and Simulation, Vol. 3 (Forschungszentrum Jülich, 2013)
http://www.cond-mat.de/events/correl13