

# Lecture notes for FYS–KJM 4480

Quantum mechanics for many-particle systems

Simen Kvaal

November 17, 2017

# Contents

<b>1</b>	<b>Preface</b>	<b>5</b>
1.1	Exercises 2017 . . . . .	6
<b>2</b>	<b>Fundamental formalism</b>	<b>7</b>
2.1	Many-particle systems . . . . .	8
2.1.1	Recap of classical mechanics . . . . .	8
2.1.2	Canonical quantization of many-particle systems . . . . .	9
2.1.3	Particle statistics and (anti-)symmetry . . . . .	11
2.1.4	The manybody Hamiltonian . . . . .	12
2.1.5	Slater determinants . . . . .	13
2.1.6	Exercises . . . . .	16
2.2	Second quantization . . . . .	19
2.2.1	Creation and annihilation operators . . . . .	19
2.2.2	Building the Slater determinants recursively . . . . .	21
2.2.3	Anticommutator relations . . . . .	22
2.2.4	Occupation number representation . . . . .	23
2.2.5	Fock space . . . . .	24
2.2.6	Exercises . . . . .	26
2.3	Representation of operators . . . . .	31
2.3.1	What we will prove . . . . .	31
2.3.2	One-body operators . . . . .	32
2.3.3	Two-body operators . . . . .	33
2.3.4	Diagonalizing a one-body operator . . . . .	34
2.3.5	Exercises . . . . .	34
2.4	Many-fermion systems: examples . . . . .	40
2.4.1	Electronic systems . . . . .	40
2.4.2	The $N$ -electron parabolic quantum dot in one dimension . . . . .	41
2.4.3	The nuclear manybody problem . . . . .	42
2.4.4	Exercises . . . . .	43
<b>3</b>	<b>Hartree–Fock theory part I</b>	<b>45</b>
3.1	The Hartree–Fock equations . . . . .	45
3.1.1	Exercises . . . . .	49
3.2	The direct and exchange potentials . . . . .	51
3.3	Augmenting the HF single-particle functions to a basis . . . . .	52
3.3.1	Exercises . . . . .	54
3.4	The Hartree–Fock equations in a given basis: the Roothan–Hall equations . . . . .	56

3.4.1	Exercises . . . . .	57
3.5	Self-consistent field iteration . . . . .	58
3.6	RHF: Restricted Hartree–Fock [needs work] (optional) . . . . .	59
3.7	UHF: Unrestricted Hartree–Fock theory [needs work] (optional) . . . . .	62
<b>4</b>	<b>Wick’s Theorem and normal ordering</b>	<b>63</b>
4.1	Wick’s Theorem . . . . .	63
4.1.1	A motivation . . . . .	63
4.1.2	Vacuum expectation values . . . . .	64
4.1.3	Normal ordering and contractions . . . . .	65
4.1.4	Statement of Wick’s Theorem . . . . .	67
4.1.5	Vacuum expectation values using Wick’s Theorem . . . . .	67
4.1.6	Exercises . . . . .	68
4.2	Proof of Wick’s Theorem . . . . .	69
4.2.1	A lemma . . . . .	69
4.2.2	Proof of Wick’s Theorem using the lemma . . . . .	71
4.3	Using Wick’s Theorem . . . . .	72
4.3.1	Exercises . . . . .	72
4.4	The Generalized Wick’s Theorem . . . . .	73
4.4.1	Exercises . . . . .	73
4.5	Normal ordering relative to Hartree–Fock vacuum: particle-hole formalism . . . . .	75
4.5.1	Exercises . . . . .	76
4.6	Operators on normal-order form . . . . .	78
4.6.1	The number operator . . . . .	78
4.6.2	One-body operators . . . . .	78
4.6.3	Two-body operators . . . . .	79
4.6.4	Normal-ordered two-body Hamiltonian . . . . .	80
4.6.5	Full expressions for the normal-ordered Hamiltonian . . . . .	81
<b>5</b>	<b>Hartree–Fock theory part II</b>	<b>83</b>
5.1	Normal-ordered Hamiltonian using HF basis . . . . .	83
5.2	Brillouin’s Theorem . . . . .	84
5.3	Stability of HF . . . . .	85
5.3.1	Exercises . . . . .	87
<b>6</b>	<b>The Standard Methods of approximation</b>	<b>88</b>
6.1	Introduction . . . . .	88
6.2	The variational principle (Not lectured 2017) . . . . .	89
6.2.1	The Cauchy interlace theorem and linear models . . . . .	92
6.2.2	Exercises . . . . .	93
6.3	The Configuration-interaction method (CI) . . . . .	94
6.3.1	General description . . . . .	94
6.3.2	Matrix elements of the CI method . . . . .	95
6.3.3	Computer implementation of CI methods . . . . .	96
6.3.4	Naive CI . . . . .	96
6.3.5	Direct CI . . . . .	97
6.3.6	Recipe for bit pattern representation. . . . .	98
6.3.7	Exercises . . . . .	99
6.4	Perturbation theory for the ground-state (PT) . . . . .	100

6.4.1	Non-degenerate Rayleigh–Schrödinger perturbation theory (RSPT)	100
6.4.2	Low-order RSPT	103
6.4.3	A two-state example	104
6.4.4	Manybody Perturbation Theory (MBPT)	106
6.4.5	Møller–Plesset Perturbation Theory	108
6.4.6	Exercises	110
<b>7</b>	<b>The Electron Gas</b>	<b>113</b>
7.1	The Jellium Model	113
7.1.1	Fourier series and plane-wave basis sets	113
7.1.2	Non-interacting electrons in a box: Fermi gas	115
7.1.3	Interacting gas/Jellium: Hamiltonian	116
7.1.4	Hamiltonian in second quantization	119
7.1.5	Hartree–Fock treatment	120
7.1.6	Evaluation of sums in HF energies	121
7.1.7	Spin polarization: Unrestricted Hartree–Fock	122
7.1.8	Perturbation theory for jellium is divergent	123
<b>8</b>	<b>Common basis sets</b>	<b>127</b>
8.1	Harmonic oscillator basis functions	128
8.1.1	$d$ -dimensional HO in Cartesian coordinates	128
8.1.2	Polar coordinate HO eigenfunctions, $d = 2$	130
8.1.3	Polar coordinate HO eigenfunctions, $d = 3$	131
8.2	Plane-wave basis set	132
8.3	The nuclear manybody problem in a non-relativistic approximation	133
8.3.1	The self-bound property of the nucleus, and removal of center-of-mass degree of freedom	133
8.4	Molecular systems and Gaussian basis sets	134
8.4.1	The Born–Oppenheimer molecular Hamiltonian	134
8.4.2	Hartree–Fock and Post-Hartree–Fock methods	135
8.4.3	From hydrogenic to Gaussian orbitals	136
8.4.4	Gaussian basis sets	137
8.4.5	Gaussians are useful because they give fast integration	137
<b>9</b>	<b>Coupled-cluster theory (CC)</b>	<b>138</b>
9.1	Motivation and introduction: Cluster functions and the exponential ansatz	139
9.2	Excitation operators, cluster operators, existence of cluster expansion	143
9.3	The coupled-cluster equations	146
9.3.1	Unlinked equations	146
9.3.2	Linked equations	146
9.3.3	Wick’s Theorem and The Baker–Campbell–Hausdorff expansion	147
9.3.4	Full CC versus truncated CC	148
9.4	The CCSD amplitude equations written out	149
9.4.1	A variational CC theory?	149
9.5	Numerical solution of the amplitude equations	150
9.6	The coupled-cluster Lagrangian	151
9.6.1	Expectation values	151
9.6.2	The Hellmann–Feynman theorem	151
9.6.3	CC expectation values I: direct approach	152

9.6.4	CC expectation values II: The CC Lagrangian formalism . . . . .	152
9.7	The Equation-of-motion CC method for excited states . . . . .	155
9.7.1	Exercises . . . . .	156
9.8	Coupled-cluster diagrams . . . . .	157
9.8.1	Introduction . . . . .	157
9.8.2	Diagram elements . . . . .	157
9.8.3	CC energy equation . . . . .	157
9.8.4	CCSD amplitude equations . . . . .	157
9.8.5	Exercises . . . . .	157
<b>10</b>	<b>Multireference systems</b>	<b>158</b>
10.1	Single-reference vs. multireference . . . . .	158
10.2	Multiconfiguration Hartree–Fock . . . . .	160
10.2.1	Wavefunction ansatz . . . . .	160
10.2.2	Unitary transformations of SPF bases . . . . .	161
10.2.3	Energy expression . . . . .	162
10.2.4	Variational equations: variation of $A$ . . . . .	163
10.2.5	Variational equations: variation of $\phi^{\text{act}}$ . . . . .	163
10.2.6	Newton’s Method for MCHF . . . . .	166
10.2.7	Complete active space self-consistent field (CAS-SCF) . . . . .	166
10.2.8	Exercises . . . . .	167
<b>11</b>	<b>Introduction to time-dependent theory (Not lectured 2017)</b>	<b>169</b>
<b>A</b>	<b>Mathematical supplement</b>	<b>170</b>
A.1	Calculus of variations . . . . .	171
A.1.1	Functionals . . . . .	171
A.1.2	Functions of one real variable . . . . .	171
A.1.3	Functions of two real variables . . . . .	171
A.1.4	Extremalization of a functional . . . . .	173

# Chapter 1

## Preface

This set of lecture notes has been in development since fall 2015. The course deals with the quantum formalism needed for efficiently dealing with systems of many identical particles from the nonrelativistic quantum mechanical point of view. particles.

The quantum theory of few-body systems, such as the hydrogen atom, is covered in most introductory courses on quantum mechanics. Usually, it is the perspective of *wave mechanics* which is taken, solving the Schrödinger equation as a partial differential equation (PDE). When dealing with systems of *more than two* particles, the wavefunction grows rapidly in complexity, and the language of *second quantization* becomes indispensable. The focus turns from the PDE point of view to an operator point of view – matrix mechanics. For example, it is almost always necessary to introduce *approximate calculations*, as the Schrödinger equation is usually too complex to solve exactly.

The focus of the course leans towards systems of many electrons, typically molecular systems or quantum dot models.

## 1.1 Exercises 2017

All exercises are *strongly recommended*. From Week 2, the exercises further develop material lectured, bringing in essential tools that we will use.

Week 1 (Aug 21–Aug 27):

- 2.2 on page 16
- 2.4 on page 17
- 2.7 on page 17
- 2.8 on page 18
- 2.9 on page 18

Week 2 (Aug 28–Sep 3):

- 2.18 on page 29
- 2.19 on page 29
- 2.21 on page 34
- 2.22 on page 35
- 2.23 on page 36
- 3.2 on page 49
- 3.1 on page 49
- 3.3 on page 50
- 3.4 on page 54
- 3.5 on page 54

Week 3 (Sep 4–Sep 10):

- 4.1 on page 68
- 4.2 on page 68
- 4.3 on page 68
- 4.4 on page 68
- 4.6 on page 72
- 4.7 on page 73

Week 4 (Sep 11–Sep 17): Project 1.

## Chapter 2

# Fundamental formalism

Suggested reading for this chapter: Raimes [1], sections 1.1–1.3, and Gross/Runge/Heinonen [2], section I.1. Chapter 1 of Szabo/Ostlund [3] contains a nice refresher on mathematical topics, including linear algebra.

The lecture notes by Jon Magne Leinaas from the course FYS4410 are great for the general quantum formalism.



## 2.1 Many-particle systems

### 2.1.1 Recap of classical mechanics

In order to properly introduce quantum systems of many particles, it is useful to first remind ourselves about the classical theory of many particles. The mathematical formulation of quantum mechanics resembles classical mechanics to a large extent, and the procedure known as “canonical quantization” turns any classical model into a quantum mechanical model. The material in this section will not be used later in the course, but serves as background material.

A classical (Newtonian) model system of  $N$  particles is described by (1) the *state* of the system, and (2) a governing equation of motion, specifying how the particles move. To completely specify the state of the  $N$  identical particles we need not only the coordinates  $q = (\vec{r}_1, \dots, \vec{r}_N)$ , but also the momenta  $p = (\vec{p}_1, \dots, \vec{p}_N)$ . Together, they form a point  $\xi = (q, p)$  in *phase space*  $\in \mathbb{R}^{2 \cdot n}$ . Here,  $n = Nd$ , where  $d$  is the spatial dimension of physical space. Usually,  $d = 3$ , but for some models effectively  $d < 3$ , such as a quantum dot ( $d = 2$ ), a quantum wire ( $d = 1$ ), or a 2D electron gas.

The governing equation of motion is *Hamilton's equation of motion*,

$$\dot{q} = \frac{\partial}{\partial p} \mathcal{H}(q, p) \quad (2.1a)$$

$$\dot{p} = -\frac{\partial}{\partial q} \mathcal{H}(q, p), \quad (2.1b)$$

where  $\mathcal{H}(q, p)$  is the *Hamiltonian*, a function that is the total energy of the system.

For  $N$  identical particles of mass  $m$ , interacting among themselves with a force originating from a central potential  $w(r_{ij})$ ,  $r_{ij} = |\vec{r}_i - \vec{r}_j|$ , and moving in an external potential field  $v(\vec{r})$ , the Hamiltonian is given by

$$\begin{aligned} \mathcal{H}(q, p) &= \mathcal{T}(p) + \mathcal{V}(q) + \mathcal{W}(q) \\ &= \frac{1}{2m} \sum_{i=1}^N |\vec{p}_i|^2 + \sum_{i=1}^N v(\vec{r}_i) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N w(r_{ij}). \end{aligned} \quad (2.2)$$

The first sum is the *kinetic energy*  $\mathcal{T}$ , while the second sum is the *external potential energy*  $\mathcal{V}$ , since it does not arise from the particles themselves. The final sum is the *interaction energy*  $\mathcal{W}$ .

It is common to introduce the symplectic matrix

$$J = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix}, \quad (2.3)$$

which satisfies  $J^2 = -I_{2n}$ , i.e., it acts as an imaginary unit, and write

$$J \dot{\xi} = \frac{\partial}{\partial \xi} \mathcal{H}(\xi). \quad (2.4)$$

Suppose  $\xi(t)$  solves Eq. (2.4). Then a simple calculation shows that

$$\frac{d}{dt} E(t) = \frac{d}{dt} \mathcal{H}(\xi(t)) = \frac{\partial \mathcal{H}}{\partial \xi} \cdot J^{-1} \frac{\partial \mathcal{H}}{\partial \xi} = \frac{\partial \mathcal{H}}{\partial q} \cdot \frac{\partial \mathcal{H}}{\partial p} - \frac{\partial \mathcal{H}}{\partial p} \cdot \frac{\partial \mathcal{H}}{\partial q} = 0. \quad (2.5)$$

Indeed, let  $\omega(q, p)$  be any differentiable real-valued function, an *observable*. Then the total time derivative is given by

$$\frac{d}{dt} \omega(\xi(t)) = \frac{\partial \omega}{\partial \xi} \cdot J^{-1} \frac{\partial \mathcal{H}}{\partial \xi} := \{\omega, \mathcal{H}\}(\xi(t)), \quad (2.6)$$

where the *Poisson bracket* is defined by

$$\{f, g\} := \frac{\partial f}{\partial \xi} \cdot J^{-1} \frac{\partial g}{\partial \xi} = \frac{\partial f}{\partial q} \cdot \frac{\partial g}{\partial p} - \frac{\partial f}{\partial p} \frac{\partial g}{\partial q}. \quad (2.7)$$

The Poisson bracket is thus a map that takes two *functions*  $f$  and  $g$  over phase space into a new function  $\{f, g\}$  over phase space.

In particular, we have the brackets

$$\{q_k, p_\ell\} = \delta_{k\ell}, \quad (2.8)$$

where now  $k, \ell = 1, \dots, n$ . You are encouraged to verify this by computation. Any coordinate system  $\xi = (q, p)$  for our system that satisfies Eq. (2.8) are called *canonical*.

At this point, exercise 2.1 is recommended.

Since the particles are *identical*, it is not meaningful to put a label on the particles. It is only the *set*  $\{\vec{r}_1, \dots, \vec{r}_N\}$  which can be distinguished, not the order of the particles. This implies that the vector  $q$  of particle coordinates is only well-defined up to a permutation of the particles, i.e., we must *identify*  $q$  and a permutation  $q'$  of  $q$ . Thus, the set of  $N$  particles “live in” the quotient space  $\mathbb{R}^{Nd}/S_N$ , where  $S_N$  is the group of permutations [4]. Usually we don’t introduce this quotient space when working with Hamilton’s equations of motion, as they will be the same. Instead we implement the invariance via the *observables*. A classical observable is any real-valued function of the state of the system, i.e., of the phase space point  $\xi$ . Since the phase space has permutations factored out, this means that  $\omega$  is an observable if and only if  $\omega(\xi) = \omega(\xi')$  for any permutation  $\xi'$  of the phase-space point  $\xi$ .

As an example, a particle’s individual position  $\vec{r}_i$  is not an observable, since it is not permutationally invariant. Indeed, to say “individual particle” is logically inconsistent! But the center of mass  $\vec{R} = (\sum_i \vec{r}_i)/N$  is invariant, and thus an observable.

## 2.1.2 Canonical quantization of many-particle systems

### States: From phase space to Hilbert space

It is amusing that Equation (2.4) has a striking resemblance to the Schrödinger equation, especially since  $J^2 = -I_{2n}$  is like an imaginary unit. Indeed, via a procedure known as *canonical quantization* we obtain from the classical Hamiltonian description the corresponding non-relativistic quantum mechanical description of the particles.

It is a fact that identical particles like protons, neutrons, and electrons have extra internal degrees of freedom that do not show up classically, i.e., *spin*. For an electron, the spin is a vector that can take on the lengths  $\pm\hbar/2$  along some arbitrary axis. Each electron therefore has a coordinate  $x = (\vec{r}, s)$ , where  $s = \pm\hbar/2$  is the spin projection along this axis. A nucleon has an additional *isospin*  $\tau = \pm\hbar/2$  that “chooses” whether the nucleon is a proton or a neutron. We let  $X$  be the configuration space of a single particle, i.e.,  $x \in X$ . Note that the introduction of these additional degrees of freedom is *ad hoc* in this section – they cannot be derived from the described quantization procedure. Instead, they are consequences of other physical arguments, e.g., Dirac predicted spin using relativistic arguments.

Whereas the classical state of such a system is a point in phase space, the quantum state is a complex-valued wavefunction in an infinite dimensional Hilbert space, i.e., a complete vector space with an inner product. This Hilbert space can be chosen in a variety of ways, but the standard choice (see paragraph below on observables) is the space of square-integrable functions  $\Psi$  depending on all the coordinates:

$$\Psi = \Psi(x_1, x_2, \dots, x_N) \quad (2.9)$$

Here,  $(x_1, \dots, x_N) \in X^N$  is a point in the configuration space of  $N$  particles. Thus, the wavefunction is a map

$$\Psi : X^N \longrightarrow \mathbb{C}. \quad (2.10)$$

The wavefunction has a probabilistic interpretation:  $P(x_1, \dots, x_N) = |\Psi(x_1, x_2, \dots, x_N)|^2$  is the probability density for locating all particles at the point  $(x_1, \dots, x_N) \in X^N$ . Since the total probability must be 1,  $\Psi$  must be *square integrable*, i.e.,

$$1 = \int_{X^N} |\Psi(x_1, x_2, \dots, x_N)|^2 dx_1 dx_2 \dots dx_N \equiv \sum_{s_1, s_2, \dots, s_N} \int |\Psi(\vec{r}_1, s_1, \vec{r}_2, s_2, \dots, \vec{r}_N, s_N)|^2 d\vec{r}_1 \dots d\vec{r}_N. \quad (2.11)$$

On the other hand, the contribution to the integral over sets of measure zero (such as individual points, or lines, surfaces, ...) is zero. Thus,  $\Psi$  is only well defined up to its values on a set of measure zero.  $\Psi'$  and  $\Psi$  describe the *same probability density* if

$$\int_{X^N} |\Psi' - \Psi|^2 dx_1 \dots dx_N = 0. \quad (2.12)$$

Mathematically, the set of square integrable functions that differ from one another on at most a set of measure zero constitute an element in a Hilbert space called  $L^2(X^N)$ . The elements of  $L^2(X^N)$  are *equivalence classes*  $[\Psi]$  consisting of all the functions  $\Psi' : X^N \rightarrow \mathbb{C}$  such that  $\Psi = \Psi'$  everywhere, except for possibly on a set of measure zero. This is a mathematical subtlety that can resolve some apparent dilemmas and paradoxes, but that we rarely encounter in practice. Usually, even mathematicians simply say that  $\Psi \in L^2(X^N)$ , even if it is not strictly true.

In classical mechanics, we also needed momenta to specify the state of the system. In quantum mechanics, the momenta are replaced by operators, see below, and the fact that  $\Psi$  is complex-valued.

### Observables: from real-valued functions to self-adjoint operators

It is a postulate of quantum mechanics that *all physics* can be obtained from the state  $\Psi$  via “expectation values of observables”. In classical mechanics, an observable was a (permutation invariant) function  $\omega$  of the state  $\xi \in \mathbb{R}^{2n}$ . The *value* of the observable was  $\omega(\xi)$  for a system in the state  $\xi$ . In quantum mechanics, an observable is a *self-adjoint operator* (Hermitian operator)  $\hat{\Omega}$ , and its *value* in the *state*  $\Psi$  is the “expectation value”

$$\langle \Omega \rangle = \langle \Psi | \hat{\Omega} | \Psi \rangle = \int \Psi(x_1, \dots, x_N)^* [\hat{\Omega} \Psi(x_1, \dots, x_N)] dx_1 \dots dx_N. \quad (2.13)$$

How do we obtain quantum observables from the classical observables? The basic rule of canonical quantization is to replace a set of classically canonical coordinates  $\xi = (q, p)$  with operators  $(\hat{q}, \hat{p})$  such that

$$\{q_j, p_k\} = \delta_{jk} \implies [\hat{q}_j, \hat{p}_k] = i\hbar \delta_{jk}. \quad (2.14)$$

This does not *uniquely* determine  $(\hat{q}, \hat{p})$  but one can show that the different solutions are equivalent (in the sense that they can be put in a unitary correspondence). The Hilbert space is also part of the solution. We have chosen the space  $L^2(X^N)$ , and therefore  $\hat{q}_j$  must be the operator that multiplies with  $q_j$ , i.e.,  $\hat{r}_j$  is the operator that multiplies with  $\vec{r}_j$ :

$$[\hat{r}_i \Psi](x_1, \dots, x_N) = \vec{r}_i \Psi(x_1, \dots, x_N). \quad (2.15)$$

Given the choice of solution for  $\hat{q}_j$ , the standard choice for  $\hat{p}_j$  is

$$\hat{p}_j = -i\hbar \frac{\partial}{\partial q_j}, \quad (2.16)$$

i.e.,  $\hat{p}_i = -i\hbar \nabla_i$ :

$$[\hat{p}_i \Psi](x_1, \dots, x_N) = -i\hbar \nabla_i \Psi(x_1, \dots, x_N). \quad (2.17)$$

Given a classical observable  $\omega(q, p)$ , the quantum observable is given by

$$\hat{\Omega} = \omega(\hat{q}, \hat{p}). \quad (2.18)$$

(This sometimes leads to non-Hermitian observables. In that case one needs to symmetrize  $\hat{\Omega}$  properly. We will not discuss this potential problem.)

In particular, the Hamiltonian function  $\mathcal{H}(q, p)$  is mapped into the Hamiltonian operator  $\hat{H}$ . The Hamiltonian operator corresponding to Eq. (2.2) is discussed below.

We have elegantly swept several mathematical details under the rug. For the mathematical well-definedness of the theory, it is essential that the observables are *self-adjoint* operators. This is a criterion that is stronger than merely being Hermitian. We will not go into the details here, as they will not matter for this course.

### Time evolution: from Hamilton's equation to the Schrödinger equation

The governing equation in non-relativistic quantum mechanics is the time-dependent Schrödinger equation (TDSE):

$$\hat{H}\Psi(x_1, x_2, \dots, x_N, t) = i\hbar \frac{\partial}{\partial t} \Psi(x_1, x_2, \dots, x_N, t). \quad (2.19)$$

The TDSE has a striking resemblance to Hamilton's equations of motion for the classical counterpart. However, the first order *ordinary* differential equation is replaced by a first-order *partial* differential equation.

We have assumed in this discussion that  $\hat{H}$  (or  $\mathcal{H}$ ) does not depend on time. Under that assumption, we have a formal solution of the TDSE,

$$\Psi(x_1, \dots, x_N; t) = U(t)\Psi(x_1, \dots, x_N, 0) = e^{-it\hat{H}/\hbar}\Psi(x_1, \dots, x_N). \quad (2.20)$$

It is well-known that if we can *diagonalize*  $\hat{H}$ , the TDSE is “solved”. Moreover, a quantum system invariably coupled to its environment, leading to a relaxation towards the *ground state*. We will therefore focus on the time-independent Schrödinger equation in this course:

$$\hat{H}\Psi(x_1, x_2, \dots, x_N) = E\Psi(x_1, x_2, \dots, x_N). \quad (2.21)$$

The time-independent Schrödinger equation is the main focus in this course, and we will only scratch the surface.  $\Psi$  is a very, very, *very* complicated function. Intuitively, one might think that solving for  $\Psi$  is  $N$  times as hard as solving for an  $N = 1$  wavefunction. However,  $\Psi$  is a function of *all*  $N$  coordinates. Resolving each coordinate on a grid with, say,  $K$  points requires  $K^N$  points in total. For  $K = 2$  (which is rather coarse) and  $N = 40$  (e.g., a  $^{40}\text{Ca}$  nucleus), we need  $2^{40} \approx 10^{12}$  data points! Describing the correlated motion of  $N$  quantum particles is harder than the pioneers of quantum mechanics thought! Literally *thousands* of researchers worldwide are make a living out of devising more or less clever schemes for finding approximate solutions.

### 2.1.3 Particle statistics and (anti-)symmetry

Our particles are identical, or indistinguishable. There is abundant evidence that all elementary particles must be treated as such. That means that our probability density must be *permutation invariant* in the following sense: let  $\sigma \in S_N$  be a permutation of  $N$  indices, and let  $(x_1, \dots, x_N) \in X^N$  be a configuration of the  $N$  particles. Then we must have

$$|\Psi(x_1, x_2, \dots, x_N)|^2 = |\Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)})|^2. \quad (2.22)$$

This is equivalent to

$$\Psi(x_1, \dots, x_N) = e^{i\alpha(\sigma)} \Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)}) \quad (2.23)$$

for some real  $\alpha(\sigma)$ , that may depend on  $\sigma$ .

Define a linear operator  $\hat{P}_\sigma$  via

$$(\hat{P}_\sigma \Psi)(x_1, \dots, x_N) = \Psi(x_{\sigma(1)}, x_{\sigma(2)}, \dots, x_{\sigma(N)}), \quad (2.24)$$

that is, the operator that evaluates  $\Psi$  at permuted coordinates. We have reformulated particle indistinguishability as:  $\Psi$  is an eigenfunction of  $\hat{P}_\sigma$  for every  $\sigma \in S_N$ , with eigenvalue possibly depending on  $\sigma$ .

One can show (see the exercises), that either  $\hat{P}_\sigma \Psi = \Psi$  for every  $\sigma \in S_N$ , or  $\hat{P}_\sigma \Psi = (-1)^{|\sigma|} \Psi$  for every  $\sigma \in S_N$ , where  $|\sigma|$  is the number of transpositions in  $\sigma$ , and thus  $(-1)^{|\sigma|}$  is the sign of the permutation. In the former case,  $\Psi$  is “totally symmetric with respect to permutations”, and in the latter case, “totally anti-symmetric”.

It is a *postulate* that particles occurring in quantum theory (in three-dimensional space) are of one of two types: bosons or fermions. Bosons have totally symmetric wavefunctions only, and fermions have totally anti-symmetric wavefunctions only. To cite Leinaas and Myrheim [4], “The physical consequences of this postulate seem to be in good agreement with experimental data.” Wolfgang Pauli proved (using relativistic considerations) that wavefunctions of half-integral *spin* must be anti-symmetric, and wavefunctions of particles with integral spin must be symmetric, connecting the postulate with the intrinsic spin of particles. To this day, no particles with other spin values have been found.

In this course, we focus on fermions. See, e.g., [2] for the general case.

As a side note, it was the Norwegian physicists Jon Magne Leinaas and Jan Myrheim [4] who first did a careful analysis of the configuration space of  $N$  identical classical particles, and quantizing this system in a proper fashion. They showed how the fermionic and bosonic symmetries were *consequences of the structure of the classical configuration space*. They also discovered that in 1D or 2D, more general permutation symmetries are possible, leading to a third type of fundamental particles, “anyons”.

### 2.1.4 The manybody Hamiltonian

We consider the manybody Hamiltonian in some detail, and as it covers most of our needs, let us quantize the classical Hamiltonian (2.2), assuming that all the particles present are electrons with mass  $m_e$ .

The first sum is the kinetic energy  $\mathcal{T}(\vec{p}_1, \dots, \vec{p}_N)$ . The kinetic energy of a single electron is  $|\vec{p}|^2/(2m_e)$ . The corresponding operator is

$$\hat{t} = \frac{1}{2m_e} |\hat{p}|^2 = \frac{1}{2m_e} (-i\hbar \nabla)^2 = -\frac{\hbar^2}{2m_e} \nabla^2. \quad (2.25)$$

Summing over particles, we get

$$\hat{T} := \mathcal{T}(\hat{p}_1, \dots, \hat{p}_N) = \sum_{i=1}^N \hat{t}(i) = -\frac{\hbar^2}{2m_e} \sum_{i=1}^N \nabla_i^2. \quad (2.26)$$

Now, let us consider the external potential part of Eq. (2.2),  $\mathcal{V}(\vec{r}_1, \dots, \vec{r}_N)$ . This becomes

$$\hat{V} := \mathcal{V}(\hat{r}_1, \dots, \hat{r}_N) = \sum_{i=1}^N v(\hat{r}_i), \quad (2.27)$$

a multiplication operator.

For a Coulomb field set up by an atomic nucleus of charge  $+Ze$  at the location  $\vec{R}$ , we have

$$v(\vec{r}) = -\frac{Ze^2}{\|\vec{R} - \vec{r}\|}. \quad (2.28)$$

The electron pair  $(i, j)$  interacts via the Coulomb force:

$$w(i, j) = \frac{e^2}{|\vec{r}_i - \vec{r}_j|}. \quad (2.29)$$

Thus,

$$\hat{W} = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N w(i, j) = \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \frac{e^2}{|\hat{r}_i - \hat{r}_j|}. \quad (2.30)$$

The complete Hamiltonian becomes

$$\hat{H} = \hat{T} + \hat{V} + \hat{W} = \hat{H}_0 + \hat{W}, \quad (2.31)$$

where  $\hat{H}_0$  is a *one-body operator* since it contains terms that only involve one particle at a time. The operator  $\hat{W}$  is a *two-body operator*.

In this course, the typical generic form of the quantum Hamiltonian is:

$$\begin{aligned} \hat{H} &= \sum_{i=1}^N \hat{h}(i) + \frac{1}{2} \sum_{\substack{i,j=1 \\ i \neq j}}^N \hat{w}(i, j) \\ &= \hat{H}_0 + \hat{W}. \end{aligned} \quad (2.32)$$

where  $\hat{h}(i)$  denotes a single-particle operator acting only on the degrees of freedom of particle  $i$ , and  $\hat{w}(i, j) = \hat{w}(j, i)$  denotes a two-body operator that acts only on the degrees of freedom of the *pair*  $(i, j)$ ,  $i \neq j$ .

Of course, one could consider three-body forces as well, and even higher. Such often occur in nuclear physics, where the non-relativistic Hamiltonian arises from quantum-field theoretical approximations. We will rarely have occasion to work with such operators in this course.

### 2.1.5 Slater determinants

We have established that the wavefunction of  $N$  identical fermions must be totally antisymmetric and square integrable. The set of such functions form a Hilbert space is denoted by  $L_N^2$  in this course.

The antisymmetry of a wavefunction of  $N$  coordinates is a quite complicated constraint. It is a formidable task to solve the Schrödinger equation on PDE form and implement these constraints. On the other hand, if we had an orthonormal *basis* with such constraints, we could express the PDE problem as an (infinite dimensional) matrix problem instead.

Consider the Hilbert space of a *single* particle,  $L^2(X)$ , consisting of *single-particle functions* (SPFs). Suppose we are given an orthonormal SPF basis  $\{\phi_1, \phi_2, \dots, \phi_p, \dots\}$ , with  $\langle \phi_p | \phi_q \rangle = \delta_{pq}$ , such that any wavefunction can be written

$$\psi(x) = \sum_p \psi_p \phi_p(x), \quad \psi_p = \langle \phi_p | \psi \rangle. \quad (2.33)$$

Taking the *tensor product* of  $N$  single-particle functions, we get a square-integrable function of  $N$  particles,

$$\tilde{\Phi}_{p_1 \dots p_N}(x_1, \dots, x_N) = \phi_{p_1}(x_1) \phi_{p_2}(x_2) \dots \phi_{p_N}(x_N). \quad (2.34)$$

The collection of functions  $\tilde{\Phi}_{p_1, \dots, p_N}$  forms an orthonormal basis for the square-integrable functions of  $N$  particles. However, they are *not antisymmetric*, and thus not an orthonormal basis for  $L_N^2$ .

We can construct an antisymmetric basis using *Slater determinants*, which we now describe.

What is the simplest totally antisymmetric wavefunction we can create, starting with some given single-particle functions? If we start with  $N = 2$ , and consider the product  $\phi_1(x_1)\phi_2(x_2)$ , this is not anti-symmetric. But if we consider the linear combination

$$\Phi(x_1, x_2) = \phi_1(x_1)\phi_2(x_2) - \phi_2(x_1)\phi_1(x_2), \quad (2.35)$$

this is antisymmetric if we exchange  $x_1$  and  $x_2$ . Continuing with  $N = 3$ , we quickly realize that in order to obtain something antisymmetric out of  $\phi_1(x_1)\phi_2(x_2)\phi_3(x_3)$ , we must take the linear combination

$$\begin{aligned} \Phi(x_1, x_2, x_3) = & \phi_1(x_1)\phi_2(x_2)\phi_3(x_3) - \phi_2(x_1)\phi_1(x_2)\phi_3(x_3) - \phi_1(x_1)\phi_3(x_2)\phi_2(x_3) - \\ & \phi_3(x_1)\phi_2(x_2)\phi_1(x_3) + \phi_2(x_1)\phi_3(x_2)\phi_1(x_3) + \phi_3(x_1)\phi_1(x_2)\phi_2(x_3), \end{aligned} \quad (2.36)$$

each term representing a permutation of the indices (123). There is nothing special about (123) of course, the tuple  $(p_1 p_2 p_3)$  also works. Note already that if one of the indices are equal, then the whole linear combination is zero.

The generalization to  $N$  indices is in fact a *determinant*, and we make a definition:

**Definition 2.1.** Let  $\phi_1, \phi_2, \dots, \phi_N$  be arbitrary single-particle functions in  $L^2(X)$  (not necessarily orthonormal). The Slater determinant defined by these functions is denoted by  $[\phi_1 \phi_2 \dots \phi_N]$ , and is defined via the formula

$$\begin{aligned} [\phi_1, \phi_2, \dots, \phi_N](x_1, \dots, x_N) &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_1(x_2) & \dots & \phi_1(x_N) \\ \phi_2(x_1) & \phi_2(x_2) & \dots & \phi_2(x_N) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N(x_1) & \phi_N(x_2) & \dots & \phi_N(x_N) \end{vmatrix} \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \prod_{i=1}^N \phi_{\sigma(i)}(x_i) \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \prod_{i=1}^N \phi_i(x_{\sigma(i)}) \end{aligned} \quad (2.37)$$

Note: the  $1/\sqrt{N!}$  is there for normalization purposes, see later.

Exercise 2.6 is recommended here.

We make two simple observations: Basic properties of the matrix determinant imply that permutation of particle coordinate indices gives a sign change. Similarly, permutation of function indices gives sign change:

$$[\phi_1, \dots, \phi_i, \dots, \phi_j, \dots, \phi_N] = -[\phi_1, \dots, \phi_j, \dots, \phi_i, \dots, \phi_N] \quad (2.38)$$

$$[\phi_1, \dots, \phi_N](x_1, \dots, x_i, \dots, x_j, \dots, x_N) = -[\phi_1, \dots, \phi_N](x_1, \dots, x_j, \dots, x_i, \dots, x_N). \quad (2.39)$$

Moreover, two equal rows (i.e., equal function indices) means that two of the single-particle functions are identical, giving a vanishing determinant. If two *columns* in Eq. (2.37) are identical, the determinant vanishes. Two columns equal mean that we evaluate at some  $x_i = x_j$ . *This is the Pauli exclusion principle.*

**Theorem 2.1.** Let  $\{\phi_p\}$  be an orthonormal SPF basis for  $L^2(X)$ . Then, any  $\Psi \in L_N^2$  can be expanded in the Slater determinants

$$[\phi_{p_1}, \phi_{p_2}, \dots, \phi_{p_N}]. \quad (2.40)$$

Moreover, if we choose an ordering of the single-particle indices  $p$ , the Slater determinants with indices satisfying  $p_1 < p_2 < \dots < p_N$  form an orthonormal basis.

*Proof.* Step 1: Expand  $\Psi$  in the tensor product basis.

$$\Psi(x_1, \dots, x_N) = \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_N} \tilde{\Phi}_{p_1, \dots, p_N}(x_1, \dots, x_N). \quad (2.41)$$

Step 2: Show that the coefficients  $c_{\vec{p}}$  are antisymmetric under permutation. For simplicity, consider a transposition of  $i$  with  $j$ ,  $i < j$ :

$$\begin{aligned} \hat{P}_{ij} \Psi(x_1, \dots, x_N) &= \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_N} \hat{P}_{ij} \tilde{\Phi}_{p_1, \dots, p_N}(x_1, \dots, x_i, \dots, x_j, \dots, x_N) \\ &= \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_N} \tilde{\Phi}_{p_1, \dots, p_N}(x_1, \dots, x_j, \dots, x_i, \dots, x_N) \\ &= \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_N} \tilde{\Phi}_{p_1, \dots, p_j, \dots, p_i, \dots, p_N}(x_1, \dots, x_N) \\ &= \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_j, \dots, p_i, \dots, p_N} \tilde{\Phi}_{p_1, \dots, p_N}(x_1, \dots, x_N) \\ &= - \sum_{p_1, \dots, p_N} c_{p_1, \dots, p_N} \tilde{\Phi}_{p_1, \dots, p_N}(x_1, \dots, x_N) \end{aligned} \quad (2.42)$$

Projecting the two last inequalities onto  $\tilde{\Phi}_{q_1, \dots, q_N}$  gives

$$c_{q_1, \dots, q_j, \dots, q_i, \dots, q_N} = -c_{q_1, \dots, q_i, \dots, q_j, \dots, q_N}. \quad (2.43)$$

We decompose an arbitrary  $\sigma \in S_N$  into transpositions, and obtain

$$c_{p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(N)}} = (-1)^{|\sigma|} c_{p_1, \dots, p_N}. \quad (2.44)$$

Step 3: Rearrange summation so that we exhibit  $\Psi$  as a linear combination of Slater determinants.

Note that we can write

$$\sum_{p_1, \dots, p_N} f(p_1, \dots, p_N) = \sum_{p_1, p_2 < \dots < p_N} \sum_{\sigma \in S_N} f(p_{\sigma(1)}, \dots, p_{\sigma(N)}), \quad (2.45)$$

splitting the summation over *ordered* multiindices and permutations of these. We now get

$$\begin{aligned} \Psi &= \sum_{p_1 < \dots < p_N} \sum_{\sigma} (-1)^{|\sigma|} c_{p_1, \dots, p_N} \tilde{\Phi}_{p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(N)}} \\ &= \sum_{p_1 < \dots < p_N} (\sqrt{N!} c_{p_1, \dots, p_N}) \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \tilde{\Phi}_{p_{\sigma(1)}, p_{\sigma(2)}, \dots, p_{\sigma(N)}} \\ &= \sum_{p_1 < \dots < p_N} (\sqrt{N!} c_{p_1, \dots, p_N}) [\phi_{p_1}, \dots, \phi_{p_N}]. \end{aligned} \quad (2.46)$$

This in fact proves that the Slater determinants, when we only use ordered indices, are sufficient to expand any  $\Psi \in L_N^2$ . Clearly, if we omit one such Slater determinant, not all  $\Psi$  can be expanded. (In particular, this omitted Slater determinant cannot be expanded in the rest!) Thus, the Slater determinants with ordered indices form a basis. For the orthonormality of the Slater determinants, see Exercise 2.8  $\square$

The Slater determinant expansion is a *central tool* in many-body calculations. In practical situations, since computer memory is limited, we use a *finite* basis,  $\{\phi_1, \dots, \phi_L\}$ , generating in total  $L!/N!(L-N)!$  Slater determinants that we approximate our wavefunction with.



## 2.1.6 Exercises

**Exercise 2.1.** Suppose  $X = (Q, P) = F(\xi) = F(q, p)$  is a differentiable coordinate change such that  $F^{-1}$  exists and is differentiable, and such that

$$\{Q_j, P_k\} = \delta_{jk}.$$

Such a map is called *symplectic* or, historically, a *canonical transformation*. Show that Hamilton's equations are still valid with the energy function  $\mathcal{K}(X) = \mathcal{H}(F^{-1}(X))$ , i.e.,

$$\dot{Q} = \frac{\partial}{\partial P} \mathcal{K}(Q, P), \quad \dot{P} = -\frac{\partial}{\partial Q} \mathcal{K}(Q, P).$$

Alternatively, using symplectic notation,

$$J\dot{X} = \frac{\partial}{\partial X} \mathcal{K}(X).$$

△

**Exercise 2.2.** In this exercise, we prove that if  $\Psi \in L^2(X^N)$  is an eigenfunction for all  $\hat{P}_\sigma$ , then the eigenvalue is either 1 or  $(-1)^{|\sigma|}$ . Here,  $|\sigma|$  is the number of transpositions in an arbitrary factorization of  $\sigma$ .

[A permutation  $\tau \in S_N$  is a transposition if it exchanges only a single pair  $(i, j)$ ,  $i \neq j$ . Any  $\sigma$  can be written as a product of an either even or odd number of transpositions,

$$\sigma = \tau_{i_1 j_1} \tau_{i_2 j_2} \cdots \tau_{i_n j_n}. \quad (2.47)$$

The number  $n = |\sigma|$  is not unique, but it is always either even or odd.]

Write  $\hat{P}_{ij} \equiv \hat{P}_\tau$ .

Assume that an arbitrary (i.e., no assumptions on symmetry)  $\Psi \in L^2(X^N)$  is such that, for all  $\sigma \in S_N$ ,

$$\hat{P}_\sigma \Psi = s_\sigma \Psi.$$

Show that  $\hat{P}_{ij}^2 = 1$ , and find all the possible eigenvalues of  $\hat{P}_{ij}$ .

Under the assumption on  $\Psi$ , show that if  $s_{ij}$  is the eigenvalue of  $\hat{P}_{ij}$ ,

$$\hat{P}_{ij} \Psi = s_{ij} \Psi,$$

then, for any other pair  $(i', j')$ , the eigenvalue is  $s_{ij} = s_{i'j'}$ . You will probably need to use the group theoretical properties of permutations. (Hint: factorize  $\tau_{ij}$  into three transpositions involving a fixed arbitrary index, say 1, i.e., transpositions  $\tau_{1k}$ .)

We have established that the eigenvalue of a transposition is a characteristic of  $\Psi$ , let  $s = s_{ij}$ . Compute the eigenvalue of  $P_\sigma$  for arbitrary  $\sigma$  in terms of  $s$ . △

**Exercise 2.3.** Let

$$\hat{H} = \sum_{i=1}^N \hat{h}(i) + \sum_{(i,j)} \hat{w}(i,j).$$

Show that  $\hat{H}$  commutes with  $P_\sigma$  for any permutation  $\sigma \in S_N$ , i.e., show that for *any* wavefunction  $\Psi \in L^2(X^N)$ ,

$$\hat{H}P_\sigma\Psi = P_\sigma\hat{H}\Psi. \quad (2.48)$$

△

**Exercise 2.4.** In this exercise, we consider  $X = \mathbb{R}^3$ , i.e., no spin. Consider each of the below functions.

1.  $\Psi(\vec{r}_1, \vec{r}_2) = e^{-\alpha|\vec{r}_1 - \vec{r}_2|}$ .
2.  $\Psi(\vec{r}_1, \vec{r}_2) = \sin(\vec{e}_z \cdot (\vec{r}_1 - \vec{r}_2))$ , where  $\vec{e}_z$  is the unit vector in the  $z$ -direction.
3.  $\Psi(\vec{r}_1, \vec{r}_2, \vec{r}_3) = \sin[\vec{r}_1 \cdot (\vec{r}_2 \times \vec{r}_3)]e^{-|\vec{r}_1|^2}e^{-|\vec{r}_2|^2}e^{-|\vec{r}_3|^2}$

Answer the following questions, per function:

- Is the function totally symmetric with respect to particle permutations?
- Is the function totally antisymmetric with respect to particle permutations?
- Is the function square integrable?

△

**Exercise 2.5.** In this exercise, we define the antisymmetrization operator  $\mathcal{A}$  as

$$\mathcal{A} = \frac{1}{N!} \sum_{\sigma \in S_N} (-1)^{|\sigma|} \hat{P}_\sigma. \quad (2.49)$$

Now,

$$[\phi_1, \dots, \phi_N] = \sqrt{N!} \mathcal{A} \phi_1(x_1) \cdots \phi_N(x_N). \quad (2.50)$$

An operator  $U$  is an orthogonal projector if and only if  $U^2 = U$  and  $U^\dagger = U$ .

Prove that  $\mathcal{A}$  is an orthogonal projector from  $L^2$  onto  $L_N^2$ .

△

**Exercise 2.6.** Show that the two last lines in Eq. (2.37) are equivalent. This requires some manipulation of permutations.

△

**Exercise 2.7.** Let  $A$  be an  $N \times N$  matrix. Let  $\phi_j$ ,  $j = 1, \dots, N$  be given single-particle functions, and let  $\psi_k$ ,  $k = 1, \dots, N$  be defined by

$$\psi_k = \sum_j \phi_j A_{jk}. \quad (2.51)$$

Prove that

$$[\psi_1, \psi_2, \dots, \psi_N] = \det(A)[\phi_1, \phi_2, \dots, \phi_N]. \quad (2.52)$$

△

**Exercise 2.8.** Suppose that  $\{\phi_p\}$ ,  $p = 1, 2, \dots$  are orthonormal. Prove that  $\Phi_{p_1, \dots, p_N} = [\phi_{p_1} \phi_{p_2}, \dots, \phi_{p_N}]$  is normalized,

$$\langle \Phi_{p_1 p_2 \dots p_N} | \Phi_{p_1 p_2 \dots p_N} \rangle = 1.$$

Prove that

$$\langle \Phi_{p_1 p_2 \dots p_N} | \Phi_{q_1 q_2 \dots q_N} \rangle = 0,$$

if the indices  $q_j$  are not a permutation of the  $p_i$ . Show that if the  $q_j$  are a permutation of the  $p_i$ , i.e.,  $p_i = q_{\tau(i)}$ , then

$$\langle \Phi_{p_1 p_2 \dots p_N} | \Phi_{q_1 q_2 \dots q_N} \rangle = (-1)^{|\tau|}.$$

Show that

$$\langle \Phi_{p_1 p_2 \dots p_N} | \Phi_{q_1 q_2 \dots q_N} \rangle = \delta_{p_1 q_1} \dots \delta_{p_N q_N}.$$

if we assume that the  $p_i$  are sorted in increasing order, and the same with  $q_j$ . △

**Exercise 2.9.** This exercise tests your understanding of Slater determinants. You are supposed to use the basic definition of a Slater determinant using a matrix determinant, i.e., you do not know about second quantization in this exercise.

In this exercise, let  $\psi_i(x)$  be arbitrary single-particle functions.

- a) Expand the Slater determinant  $[\psi_1, \psi_2, \psi_3, \psi_4](x_1, x_2, x_3, x_4)$  using the definition as a matrix determinant. How many terms are there in total?
- b) Let  $A$  be a non-singular  $N \times N$  matrix, and define

$$\phi_i(x) = \sum_j \psi_j(x) A_{ji}.$$

Show that  $[\phi_1, \dots, \phi_N] = \det(A)[\psi_1, \dots, \psi_N]$ .

- c) Is it possible for a linear combination of Slater determinants to be a Slater determinant? Give an example or a proof of your answer if negative.
- c) Let  $\phi_1, \phi_2$  arbitrary single-particle functions. Let  $\Phi = [\phi_1, \phi_2]$  and  $\Psi = [\psi_1, \psi_2]$ . Compute  $\langle \Phi | \Psi \rangle$  in terms of the matrix  $S_{ij} = \langle \phi_i | \psi_j \rangle$ .
- d) Generalize c) to  $N$ -particle determinants. △

## 2.2 Second quantization

In this section, we introduce the concept of *second quantization*, where the particles themselves are discrete quanta created and destroyed with *creation and annihilation operators*.

From now on, we use the bra/ket notation, instead of explicit spatial functions. Using the bra/ket notation, quantum states are written in a basis independent way. Previously, the Slater determinant  $[\dots]$  was a specific function of the space/spin coordinates, i.e., the quantum state was written in the space/spin basis. If we, say, Fourier transform this function, we will get a *different* functional expression, representing the *same* physical state. Similarly, expanding the wavefunction in a particular basis will give rise to yet another representation – the basis coefficients. The ket notation unifies all these, and represents the abstract Hilbert space vector independent of basis. From now on, therefore  $|\Psi\rangle$  is the state, and

$$\langle x_1, x_2, \dots, x_N | \Psi \rangle \equiv \Psi(x_1, x_2, \dots, x_N). \quad (2.53)$$

Similarly, a Slater determinant is usually written as  $|\phi_{p_1} \phi_{p_2} \dots \phi_{p_N}\rangle$ ,

$$[\phi_{p_1}, \dots, \phi_{p_N}](x_1, \dots, x_N) = \langle x_1 \dots x_N | \phi_{p_1} \phi_{p_2} \dots \phi_{p_N} \rangle \quad (2.54)$$

Moreover, it is often the case that the SPF basis is understood from context, so we can drop the  $\phi$ s, i.e., use the notation  $|p_1 p_2 \dots p_N\rangle$  for the Slater determinant.

### 2.2.1 Creation and annihilation operators

In this section, we let a SPF basis be given. Consider an arbitrary Slater determinant  $|p_1 p_2 \dots p_N\rangle$  built from the SPF basis.

Recall that a basis for  $L_N^2(X^N)$  could be formed from the Slater determinants by choosing an (arbitrary) ordering of the  $p$ s, i.e.,  $|p_1 p_2 \dots p_N\rangle$  with  $p_1 < p_2 < \dots < p_N$  forms a basis for  $N$ -particle space. (If we permute the index set, we get the same function with a possible sign change, so it is not an additional basis function.)

If  $p_i = p_j$  for some  $i \neq j$ , then  $|p_1 p_2 \dots p_N\rangle = 0$  is the zero vector. We recall the antisymmetry properties,

$$\hat{P}_{ij} |p_1 \dots p_i \dots p_j \dots p_N\rangle = - |p_1 \dots p_j \dots p_i \dots p_N\rangle \quad (2.55)$$

and more generally

$$\hat{P}_\sigma |p_1 \dots p_N\rangle = (-1)^{|\sigma|} |p_{\sigma(1)} \dots p_{\sigma(N)}\rangle. \quad (2.56)$$

For any  $|\Psi\rangle \in L_N^2$ , we have the basis expansion

$$|\Psi\rangle = \sum_{\vec{p}}^{\sim} |\vec{p}\rangle \langle \vec{p} | \Psi \rangle. \quad (2.57)$$

That is, the identity operator can be written

$$1 = \sum_{\vec{p}}^{\sim} |\vec{p}\rangle \langle \vec{p}|. \quad (2.58)$$

The  $\sim$  means that we sum *only over ordered sets of indices*. As we saw earlier, the coefficients  $\langle \vec{p} | \Psi \rangle$  are permutation antisymmetric.

Looking at the determinant (2.37), we see that by adding a row containing the index  $q$ , and a column with coordinate  $x_{N+1}$ , we obtain an  $N + 1$  particle Slater determinant, modulo a constant factor. The

normalized Slater determinant is

$$\langle x_1 \cdots x_{N+1} | q p_1 p_2 \cdots p_N \rangle = \frac{1}{\sqrt{(N+1)!}} \begin{vmatrix} \phi_q(x_1) & \phi_q(x_2) & \cdots & \phi_q(x_N) & \phi_q(x_{N+1}) \\ \phi_{p_1}(x_1) & \phi_{p_1}(x_2) & \cdots & \phi_{p_1}(x_N) & \phi_{p_1}(x_{N+1}) \\ \phi_{p_2}(x_1) & \phi_{p_2}(x_2) & \cdots & \phi_{p_2}(x_N) & \phi_{p_2}(x_{N+1}) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \phi_{p_N}(x_1) & \phi_{p_N}(x_2) & \cdots & \phi_{p_N}(x_N) & \phi_{p_N}(x_{N+1}) \end{vmatrix} \quad (2.59)$$

Similarly, we can remove a row and column, and obtain an  $N - 1$  particle Slater determinant.

This inspires the *creation and annihilation operators*, that map wavefunctions between different particle number spaces:

$$c_q^\dagger : L_N^2 \rightarrow L_{N+1}^2 \quad (2.60)$$

$$c_q : L_N^2 \rightarrow L_{N-1}^2 \quad (2.61)$$

The operator  $c_q^\dagger$  is called a *creation operator* and is, roughly defined, by inserting a row and column as described. The operator  $c_q$  is the Hermitian adjoint of  $c_q^\dagger$ , and it will be shown that its action on a Slater determinant corresponds to the mentioned removal of a row and column.

We define the space  $L_0^2$  – the zero particle space – as a one-dimensional space spanned by the special ket  $|-\rangle$ , the *vacuum state*. There is nothing mysterious about this, it is just a definition that will be useful later. Note that  $|-\rangle \neq 0$ , and that  $\langle - | - \rangle = 1$ .

Recall that a linear operator is fully defined when we specify its action on a basis set. This is how we define  $c_p^\dagger$  and  $c_p$  acting on a general linear combination of Slater determinants.

**Definition 2.2** (The creation operator). *For every single-particle index  $q$ , we define the creation operator  $c_q^\dagger$  acting on the vacuum state by*

$$c_q^\dagger |-\rangle = |q\rangle. \quad (2.62)$$

*Since this is a Slater determinant with a single particle, we have, of course,  $\langle x | q \rangle = \phi_q(x)$ . For an arbitrary Slater determinant with  $N > 0$ , we define the action by*

$$c_q^\dagger |p_1 \cdots p_N\rangle \equiv |q p_1 \cdots p_N\rangle. \quad (2.63)$$

We observe already that if there is a  $j$  such that  $q = p_j$ , then  $c_q^\dagger |p_1 \cdots p_N\rangle = 0$ :

$$\hat{P}_{1j} c_q^\dagger |\vec{p}\rangle = |q \vec{p}\rangle = -|q \vec{p}\rangle = 0. \quad (2.64)$$

In terms of determinant coordinate expressions as in Eq. (2.37),  $c_q^\dagger$  inserts a column on the far right with  $x_{N+1}$  and inserts a row on the top with the index  $q$ . Finally, the whole expression is renormalized.

We note: Recall that the *basis* Slater determinants were the determinants that had *ordered indices*. Assume that  $\vec{P}$  is ordered. Clearly,  $c_Q^\dagger |\vec{P}\rangle$  is either zero or equal to  $(-1)^j |P_1 P_2 \cdots P_j Q P_{j+1} \cdots P_N\rangle$ , which is a new basis determinant. Here,  $j$  is chosen such that the augmented index set is ordered.

**Definition 2.3** (The annihilation operator). *The annihilation operator  $c_q$  is the Hermitian adjoint of  $c_q^\dagger$ .*

**Theorem 2.2** (Characterization of the annihilation operator). *We have*

$$c_q |-\rangle = 0. \quad (2.65)$$

*(There are no particles in the vacuum state that we can remove.) Let  $\vec{p}$  be any multiindex, not necessarily ordered. If  $q = p_j$  for some  $j$ , we have*

$$c_q |p_1 \cdots p_N\rangle \equiv (-1)^{j-1} |p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle. \quad (2.66)$$

*Otherwise, we have  $c_q |p_1 \cdots p_N\rangle = 0$ .*

*Proof.* Let  $|\Psi\rangle = c_q |p_1 \cdots p_N\rangle$ . We may assume that  $p_1 < p_2 < \cdots < p_N$  (why?). The state is an  $N - 1$  particle state, so we expand it in our Slater basis,

$$|\Psi\rangle = \sum_{r_1 \cdots r_{N-1}} |r_1 \cdots r_{N-1}\rangle \langle r_1 \cdots r_{N-1} | c_q | p_1 \cdots p_N \rangle. \quad (2.67)$$

Note that the expansion is over ordered indices only. By definition of the Hermitian adjoint,

$$\langle r_1 \cdots r_{N-1} | c_q | p_1 \cdots p_N \rangle = \langle p_1 \cdots p_N | c_q^\dagger | r_1 \cdots r_{N-1} \rangle. \quad (2.68)$$

We note that the bracket must be a real number by definition of  $c_q^\dagger$ , and by the exercise that computes the overlap between two Slater determinants built from orthonormal SPFs. Using the definition of  $c_q^\dagger$ ,

$$\langle r_1 \cdots r_{N-1} | c_q | p_1 \cdots p_N \rangle = \langle p_1 \cdots p_N | q r_1 \cdots r_{N-1} \rangle. \quad (2.69)$$

If  $q$  is one of the  $r_j$ , the bracket is zero. To get a non-zero contribution in the expansion, we must have  $q$  different from all the  $r_j$ . Indeed, there must be a  $j$  such that  $r_{j-1} < q < r_j$ . Write

$$c_q^\dagger |r_1 \cdots r_{N-1} r_N\rangle = |q r_1 \cdots r_{N-1} r_N\rangle = (-1)^{j-1} |r_1 \cdots r_{j-1} q r_j \cdots r_{N-1}\rangle, \quad (2.70)$$

where we note that  $j$  is such that the index set  $(r_1, r_2, \dots, q, \dots, r_{N-1})$  is an *ordered*  $N$ -particle index. As we sum over the different  $r_j$ s, we thus sum over all ordered  $N$ -particle indices where  $q$  is present in (precisely) one index. If the  $p_j$ s do not include  $q$ , all the brackets are thus zero, and we obtain  $|\Psi\rangle = 0$  in that case. Assume the opposite, that that  $q = p_k$  for some  $k$ :

$$|p_1 \cdots p_N\rangle = |p_1 \cdots p_{k-1} q p_{k+1} \cdots p_N\rangle. \quad (2.71)$$

Since both sets of indices are ordered, and since

$$\langle r_1 \cdots r_{N-1} | c_q | p_1 \cdots p_N \rangle = (-1)^{j-1} \langle p_1 \cdots p_{k-1} q p_{k+1} \cdots p_N | r_1 \cdots r_{j-1} q r_j \cdots r_{N-1} \rangle, \quad (2.72)$$

we see that we must have  $k = j$ , and  $r_i = p_i$  for  $i \leq j - 1$ , and  $r_i = p_{i+1}$  for  $i \geq j$ . Thus,

$$|\Psi\rangle = (-1)^{j-1} |p_1 p_2 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle, \quad (2.73)$$

and we are done, except for treating the case  $N = 0$ , which is left as an exercise, 2.12.  $\square$

Let us see what  $c_q$  does with the coordinate determinant expression: it corresponds to moving the  $j$ th row to the top with  $j - 1$  transpositions, giving the sign factor, and then crossing out the far right column and the first row, now containing the index  $q$ . This moving of the  $j$ th row may seem like a complication compared to the creation operator, but note that for  $c_q^\dagger$  we defined its action by *inserting*  $q$  on the top. Moving  $q$  to the  $(j + 1)$ th position will induce a  $(-1)^j$ . But  $c_q$  removes a row at an in principle arbitrary location.

### 2.2.2 Building the Slater determinants recursively

We note that the SPF basis itself is given by

$$|p\rangle = c_p^\dagger |-\rangle. \quad (2.74)$$

From a single-particle state, any two-particle determinant can be build by applying  $c_{p_1}^\dagger$  to  $|p_2\rangle$ ,

$$|p_1 p_2\rangle = c_{p_1}^\dagger c_{p_2}^\dagger |-\rangle. \quad (2.75)$$

Continuing in this fashion,

$$|p_1 \cdots p_N\rangle = c_{p_1}^\dagger \cdots c_{p_N}^\dagger |-\rangle. \quad (2.76)$$

We see that the  $N + 1$  particle Slater basis can be built from the  $N$  particle basis, by applying all the  $c_p^\dagger$ . Already this makes the creation operators useful, but what makes them *very* useful is the so-called fundamental anticommutator relations, to be described next.

### 2.2.3 Anticommutator relations

Recall that the anticommutator of two operators is defined by

$$\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}. \quad (2.77)$$

In this section, we prove three important anticommutation relations, summarized as a theorem:

**Theorem 2.3** (Important anticommutator relations).

$$\{c_{q_1}^\dagger, c_{q_2}^\dagger\} = 0 \quad (2.78a)$$

$$\{c_{q_1}, c_{q_2}\} = 0 \quad (2.78b)$$

$$\{c_{q_1}, c_{q_2}^\dagger\} = \delta_{q_1, q_2}. \quad (2.78c)$$

Equation (2.78c) is called the “fundamental anticommutator”.

*Proof.* Let  $q_1, q_2$  be a two single-particle indices, and let  $N \geq 0$  be arbitrary. By the properties of determinants, it is easy to see, that for any  $|p_1 \cdots p_N\rangle \in L_N^2$ ,

$$c_{q_1}^\dagger c_{q_2}^\dagger |p_1 \cdots p_N\rangle = -c_{q_2}^\dagger c_{q_1}^\dagger |p_1 \cdots p_N\rangle. \quad (2.79)$$

Why? The right hand side is obtained by exchanging the two first rows of the determinant on the left hand side. Note that it does not matter if one of the  $q_i$  is present among the  $p_j$ .

Since this equation holds for any basis vector, we have shown that the two creation operators *anticommute*

$$\{c_{q_1}^\dagger, c_{q_2}^\dagger\} \equiv c_{q_1}^\dagger c_{q_2}^\dagger + c_{q_2}^\dagger c_{q_1}^\dagger = 0. \quad (2.80)$$

Consequently, two annihilation operators anticommute too,

$$\{c_{q_1}, c_{q_2}\} \equiv c_{q_1} c_{q_2} + c_{q_2} c_{q_1} = 0. \quad (2.81)$$

It remains to prove that

$$\{c_{q_1}, c_{q_2}^\dagger\} \equiv c_{q_1} c_{q_2}^\dagger + c_{q_2}^\dagger c_{q_1} = \delta_{q_1, q_2}. \quad (2.82)$$

Case 1:  $q_1 = q_2 = q$ . Consider the expression

$$c_q^\dagger c_q |p_1 \cdots p_N\rangle. \quad (2.83)$$

Case 1a:  $q = q_j$  for some  $j$ . We get

$$c_q^\dagger c_q |p_1 \cdots p_N\rangle = c_q^\dagger (-1)^{j-1} |p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle = (-1)^{j-1} |p_j p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle = |p_1 \cdots p_N\rangle. \quad (2.84)$$

We also get

$$c_q c_q^\dagger |p_1 \cdots p_N\rangle = c_q |p_j p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle = 0. \quad (2.85)$$

Case 1b:  $q \notin \{p_1, \dots, p_N\}$ , i.e.,  $q$  is distinct from all the  $p_j$ . In this case,  $c_q |p_1 \cdots p_N\rangle = 0$ , so

$$c_q^\dagger c_q |p_1 \cdots p_N\rangle = 0. \quad (2.86)$$

On the other hand,

$$c_q c_q^\dagger |p_1 \cdots p_N\rangle = c_q |q p_1 \cdots p_N\rangle = (-1)^0 |p_1 \cdots p_N\rangle. \quad (2.87)$$

Case 1 can be summarized as

$$\{c_q, c_q^\dagger\} = 1 \quad (2.88)$$

as desired.

Case 2:  $q_1 \neq q_2$ . Consider

$$c_{q_1}^\dagger c_{q_2} |p_1 \cdots p_N\rangle. \quad (2.89)$$

Case 2a: If  $q_1 \in \{p_1, \dots, p_N\}$  or  $q_2 \notin \{p_1, \dots, p_N\}$ , the expression vanishes. Similarly,

$$c_{q_2} c_{q_1}^\dagger |p_1 \cdots p_N\rangle = 0. \quad (2.90)$$

Case 2b:  $q_1 \notin \{p_1, \dots, p_N\}$  and  $q_2 = p_j$  for some  $j$ :

$$c_{q_1}^\dagger c_{q_2} |p_1 \cdots p_N\rangle = (-1)^{j-1} c_{q_1}^\dagger |p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle = |p_1 \cdots p_{j-1} q_1 p_{j+1} \cdots p_N\rangle, \quad (2.91)$$

i.e.,  $p_j = q_2$  is replaced by  $q_1$ . On the other hand,

$$\begin{aligned} c_{q_2} c_{q_1}^\dagger |p_1 \cdots p_N\rangle &= c_{q_2} |q_1 p_1 \cdots p_{j-1} p_j p_{j+1} \cdots p_N\rangle \\ &= (-1)^j |q_1 p_1 \cdots p_{j-1} p_{j+1} \cdots p_N\rangle = (-1) |p_1 \cdots p_{j-1} q_1 p_{j+1} \cdots p_N\rangle. \end{aligned} \quad (2.92)$$

Summing, we see that Eq. (2.78c) is proven in general.  $\square$

What makes the anticommutators so useful, is that they *encode the antisymmetry of the Slater determinants*. Any matrix element of any operator we encounter may be evaluated using these anticommutators, replacing explicit integration over antisymmetric functions involving complicated permutation sums.

## 2.2.4 Occupation number representation

Consider a given single-particle basis  $\{\phi_p\}$  and the corresponding basis of Slater determinants  $|p_1 \cdots p_N\rangle$ .

Given a set of indices  $\{p_j\}$ , we have  $N!$  rearrangements of the indices. All of the rearrangements give rise to the same Slater determinant, up to the sign of the permutation. If  $\sigma \in S_N$  is the permutation that rearranges  $(p_1, \dots, p_N)$  into  $(q_1, \dots, q_N)$ , i.e.,  $p_{\sigma(i)} = q_i$ , then

$$|p_1 \cdots p_N\rangle = (-1)^{|\sigma|} |p_{\sigma(1)} \cdots p_{\sigma(N)}\rangle = (-1)^{|\sigma|} |q_1 \cdots q_N\rangle. \quad (2.93)$$

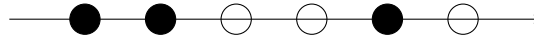
So, the *basis* of Slater determinants can be chosen as those indexed by *sorted* indices. (In some cases such a sorting of the indices is not “natural”, but it can always be made up.)

A sorted set of  $N$  indices  $(p_1, \dots, p_N)$  is in 1-1 correspondence with a *subset of integers*, or equivalently, by a picture of filled/unfilled circles, or *occupied and unoccupied sites*. One may say that the single-particle function  $\phi_{p_j}$  is *occupied* in  $|p_1 \cdots p_N\rangle$ , while any  $\phi_q$  different from all the  $p_j$  is *unoccupied*.

A common name for “single-particle function” in chemistry is “orbital”, or “spin-orbital”. We sometimes use the word “orbital” for “single-particle function”.

One can also consider a sorted index set as a *binary number* with  $N$  bits set: bit number  $q$  is set if and only if  $q \in \{p_1, \dots, p_N\}$ .

Thus, the Slater determinant  $|p_1 p_2 p_3\rangle = |014\rangle$  can be represented by the subset  $\{p_1 p_2 p_3\} = \{0, 1, 4\}$  of the set of integers, or a picture like



or the binary number

$$B = 2^{p_1} + 2^{p_2} + 2^{p_3} = 2^0 + 2^1 + 2^4 = 11001_2 = 19. \quad (2.94)$$

The different bits are called *occupation numbers*. The vacuum has no occupied single-particle functions, and is represented by the binary number 0 or the empty set.



We may use the notation

$$|n_0 n_1 \cdots n_\mu \cdots\rangle$$

to denote a Slater determinant with occupation numbers  $n_p \in \{0, 1\}$ . This Slater determinant is unique up to a sign (from permutation) and we must make sure that we, in a given context, defines the actual ordering of the SPFs precisely. For example, we choose the one determinant out of the  $N!$  possible that has  $(p_1, \dots, p_N)$  sorted:  $p_1 < p_2 < \cdots < p_N$ . In the above example,

$$|0, 1, 4\rangle = |1_0 1_1 0_2 0_3 1_4\rangle. \quad (2.95)$$

Again, we stress that occupation numbers only represent 1 of the  $N!$  Slater determinants possible to construct with  $p_1$  through  $p_N$ , namely the one where all are sorted. But they still form a basis. In the example,

$$|11001\rangle = |0, 1, 4\rangle = -|1, 0, 4\rangle = -|4, 1, 0\rangle = -|0, 4, 1\rangle = +|1, 4, 0\rangle = +|4, 0, 1\rangle, \quad (2.96)$$

exhausting all possibilities of  $N! = 3! = 6$  permutations. All these determinants are clearly linearly dependent.

The following definition can be useful:

Let  $\{p_1, \dots, p_N\}$  be an index set, and let  $q$  be an arbitrary index in  $\{p_1, \dots, p_N\}$ . Then  $\#q$  is the number of  $p_j$  that satisfies  $p_j < q$ . Thus  $\#q$  counts the occupied single-particle functions “before”  $q$ . Now, do Exercise 2.15.

## 2.2.5 Fock space

The space  $L_N^2$  has the basis consisting of the Slater determinants  $|n_0 n_1 n_2 n_3 \cdots\rangle$  with in total  $N$  occupied orbitals, or  $N$  bits set in the binary representation. The creation operator  $c_q^\dagger$  inserts a bit in position  $q$  if it is zero, possibly altering the sign of the determinant, or gives the zero vector (not the vacuum!) if bit  $q$  was already 1. Similarly, the operator  $c_q$  turns a bit off (possibly with a sign change), see Exercise 2.15

The actual number of particles in the determinant suddenly seems immaterial. Indeed, it is natural to consider *Fock space*, the direct sum of all  $L_N^2$ :

$$\mathcal{F} = \bigoplus_{N=0}^{\infty} L_N^2. \quad (2.97)$$

Taking this sum amounts to *define*  $\langle p_1 \cdots p_N | q_1 \cdots q_{N'} \rangle = 0$  whenever  $N \neq N'$ , i.e., when the occupation number expansions have different numbers of occupied sites. For example,

$$\langle 0_0 0_1 1_2 0_3 0_4 0_5 0000 \cdots | 1_0 1_1 0_2 0_3 0_4 1_5 0000 \cdots \rangle = 0, \quad (2.98)$$

since the number of particles differ in the two functions.

Now,  $c_q^\dagger : \mathcal{F} \rightarrow \mathcal{F}$  maps entirely inside  $\mathcal{F}$ , and similarly with  $c_q$ .

A basis for  $\mathcal{F}$  is the set of all  $|n_0 n_1 \cdots\rangle$  with arbitrary number of orbital occupied.

The binary number representation is quite useful for computer programs involving Slater determinants, as easily can be imagined.

In the next section, we will study the representation of many-body operators using second quantization. We make a false start by considering the *number operator*: We have that

$$c_q |n_0 n_1 \cdots\rangle = (-1)^{\#q} n_q |n_0 n_1 \cdots 0_q \cdots\rangle, \quad (2.99)$$

and furthermore that

$$c_q^\dagger c_q |n_0 n_1 \cdots\rangle = (-1)^{2\#q} n_q |n_0 n_1 \cdots\rangle. \quad (2.100)$$

Thus,

$$\sum_q c_q^\dagger c_q |n_0 n_1 \cdots\rangle = \sum_q n_q |n_0 n_1 \cdots\rangle = N |n_0 n_1 \cdots\rangle. \quad (2.101)$$

Therefore, we define

$$\hat{N} \equiv \sum_q c_q^\dagger c_q. \quad (2.102)$$

This operator extracts the number of fermions in a state  $|\Psi\rangle$  in the sense that for any  $|\Psi\rangle \in \mathcal{F}$ ,  $\hat{N}|\Psi\rangle = N|\Psi\rangle$  if and only if  $|\Psi\rangle \in L_N^2$ .

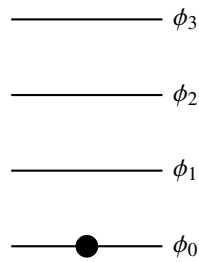
**Exercise 2.10.** [Note: This exercise has been updated since it was given as a weekly exercise.]

Let  $\phi_\mu$ ,  $\mu = 1, 2, \dots, 6$  be given orthonormal single-particle functions.

- Using the  $|\mu_1, \dots, \mu_N\rangle$  notation, write down a basis for the finite dimensional subspace of  $L^2(X^N)_{\text{AS}}$  for  $N = 2$ ,  $N = 3$  and  $N = 4$ , that you can construct using the given single-particle functions. (Make sure you include only linearly independent Slater determinants.)
- Can you construct a Slater determinant for  $N = 10$  particles using the given  $\phi_\mu$ ?
- Using the occupation number notation  $|n_1 n_2 \cdots n_6\rangle$  notation, write down a basis for the same spaces as in exercise a).
- What is the dimension of the subspace of Fock space you can create with the 6 single-particle functions?
- Assume that you have  $L$  orbitals instead of just 6. What is the dimension of the  $N$ -particle spaces you can build? What is the dimension of the Fock space you can build?

△

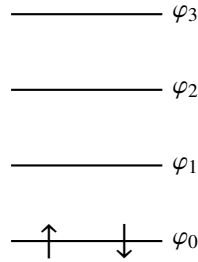
**Exercise 2.11** (Note: This exercise has been updated since it was given as a weekly exercise.). Consider the following picture:



We have four horizontal lines, each representing a single-particle function  $\phi_\mu$ . The circle represents an occupied single-particle function, i.e., the Slater determinant  $|0\rangle$ .

- In a similar fashion as the the above picture, draw a pictures of *all* the distinct Slater determinants you can create using the four single-particle functions. Make sure you consider all possible particle numbers. Caption each picture with the corresponding  $|\mu_1 \mu_2 \cdots \mu_N\rangle$ .

We now consider electrons. Consider 4 *spin-orbitals*  $\varphi_p(\vec{r})$ , i.e., 8 spin-orbitals  $\phi_\mu(\vec{r}, \sigma)$ . The corresponding diagram for the Slater determinant  $|0 \uparrow, 0 \downarrow\rangle$  is:



Each level now can hold 2 electrons, spin up and spin down.

- b) Draw all possible 2-electron Slater determinants. Mark those that have total spin projection 0.
- c) Consider the one-body operator given by

$$\hat{H}_0 = \sum_p \epsilon_p (c_{p\uparrow}^\dagger c_{p\uparrow} + c_{p\downarrow}^\dagger c_{p\downarrow}).$$

Here,  $\epsilon_p$  are numbers such that  $\epsilon_1 < \epsilon_2 < \dots$ . In first quantization,

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i).$$

Write down the matrix of the (single-electron) operator  $\hat{h}$  in the spin-orbital basis  $\{\phi_{p\sigma}\}$  and find its eigenfunctions. Interpret the spin-orbital diagram in terms of your results. Find the  $N = 4$  ground state of  $\hat{H}_0$ , and draw a picture of it.

△

## 2.2.6 Exercises

**Exercise 2.12.** Complete the proof of Theorem 2.2 in the  $N = 0$  case.

△

**Exercise 2.13.** This exercise tests your understanding of creation and annihilation operators. In this exercise,  $\phi_P$  is an orthonormal single-particle basis, and  $c_P^\dagger$  and  $c_P$  are the corresponding creation and annihilation operators, respectively. We now use the bra-ket notation for a determinant.

- a) Write down the three basic anticommutator relations. Using one of these relations, prove the following: Let  $|P_1 P_2 \dots P_N\rangle$  be a Slater determinant, and  $\sigma \in S_N$  a permutation. Then,

$$|P_1 P_2 \dots P_N\rangle = (-1)^{|\sigma|} |P_{\sigma(1)} P_{\sigma(2)} \dots P_{\sigma(N)}\rangle.$$

- b) Which of these equations are *wrong*?

1.  $c_1^\dagger c_2^\dagger c_3 |-\rangle = 0$
  2.  $c_1 c_2^\dagger c_3^\dagger |-\rangle = 0$
  3.  $c_2 c_2^\dagger c_3^\dagger |-\rangle = 0$
  4.  $c_1^\dagger c_1 c_2^\dagger |-\rangle = 0$
- c) Using the basic anticommutators, write the following as an explicit linear combination of Slater determinants. (Example:  $c_Q |P_1 P_2\rangle = c_Q c_{P_1}^\dagger c_{P_2}^\dagger |-\rangle = \delta_{Q,P_1} |P_2\rangle - \delta_{Q,P_2} |P_1\rangle$ .)
1.  $c_Q^\dagger c_P |P_1 P_2 P_3 \cdots P_N\rangle$
  2.  $c_P^\dagger c_Q c_R^\dagger c_S^\dagger |U_1 U_2\rangle$
- d) Evaluate the following expressions using the anticommutator relations as a sum of Kronecker deltas:
1.  $\langle P_1 P_2 | Q_1 \rangle$
  2.  $\langle P_1 P_2 | Q_1 Q_2 \rangle$
  3.  $\langle P_1 P_2 | c_R^\dagger c_S | Q_1 Q_2 \rangle$
  4.  $\langle - | c_{P_2} c_{P_1} c_R^\dagger c_{P_2} c_S^\dagger c_Q^\dagger | - \rangle$

△

**Exercise 2.14.** This exercise tests your understanding of occupation number formalism. We use the notation  $|n_0 n_1 \cdots n_P \cdots\rangle$  with  $n_P \in \{0, 1\}$ , for an occupation number representation of a Slater determinant built from a given set of orthonormal single-particle functions.

- a) Which of the following expressions are true? Why?
1.  $|0_0 0_1 0_2 \cdots\rangle = 0$  and/or  $|0_0 0_1 0_2 \cdots\rangle = |-\rangle$ ?
  2.  $c_1 |1_0 0_1 0_2 \cdots\rangle = |-\rangle$  and/or  $c_2 |0_0 0_1 1_2 0_3 \cdots\rangle = |-\rangle$ ?
  3.  $c_1^\dagger c_2^\dagger |-\rangle = |0_0 1_1 1_2 0_3 0_4 \cdots\rangle$  and/or  $c_1^\dagger c_2^\dagger |-\rangle = -|0_0 1_1 1_2 0_3 0_4 \cdots\rangle$ ?
- b) Write down the action of  $c_P$  and  $c_P^\dagger$  on an arbitrary  $|n_0 n_1 \cdots\rangle$ .
- c) Let  $P_1, P_2, \cdots, P_N$  be single-particle indices. Find occupation numbers  $n_P$  and a number  $s$  such that  $|P_1 \cdots P_N\rangle = (-1)^s |n_0 n_1 n_2 \cdots\rangle$ .
- d) How many fermions are there in an occupation number determinant  $|n_0 n_1 n_2 \cdots\rangle$ ?
- e) We now restrict ourselves to the case of  $L = 5$  single-particle functions.
1. Can you write down a Slater determinant with  $N = 6$  fermions?
  2. Write down a basis of Slater determinants on occupation number form for the  $N = 3$  particle Hilbert space. Also write down the *same* determinants using creation operators acting on vacuum and also using the “classical” notation  $[\phi_{P_1} \phi_{P_2} \phi_{P_3}]$ . What is the dimension of the space?
  3. What is the dimension for arbitrary  $N$ ?

4. Recall that Fock space is the space generated by all Slater determinants, with  $N$  arbitrary. What is the dimension of Fock space for  $L = 6$ ? Write down a basis on occupation number form.

△

**Exercise 2.15.** Let  $\{p_j\} = \{p_1, \dots, p_N\}$  be a given set of occupied orbitals, with occupation number representation

$$|n_0 n_1 n_2 \dots\rangle$$

Show that:

$$c_q^\dagger |n_0 n_1 n_2 \dots\rangle = \begin{cases} 0 & \text{if } q \text{ is occupied} \\ (-1)^{\#q} |n_1 n_2 \dots n_{q-1} 1_q n_{q+1} \dots\rangle & \text{if } q \text{ is unoccupied} \end{cases} \quad (2.103)$$

$$c_q |n_0 n_1 n_2 \dots\rangle = \begin{cases} 0 & \text{if } q \text{ is unoccupied} \\ (-1)^{\#q} |n_1 n_2 \dots n_{q-1} 0_q n_{q+1} \dots\rangle & \text{if } q \text{ is occupied} \end{cases} \quad (2.104)$$

△

**Exercise 2.16.** This exercise tests some basic understanding. Let  $\phi_i$ ,  $i = 1, 2, 3$  be three orthonormal single-particle functions. Consider the determinants  $|1, 2, 3\rangle$ ,  $|1, 3, 2\rangle$ ,  $|2, 1, 3\rangle$ ,  $|3, 2, 1\rangle$ ,  $|2, 3, 1\rangle$  and  $|3, 1, 2\rangle$ .

- Are there further  $N = 3$  Slater determinants that can be created using the single-particle orbitals  $\phi_i$ ,  $i = 1, 2, 3$  only?
- Write down a basis for the space spanned by the six determinants, i.e., a basis for all the vectors on the form

$$|\Psi\rangle = a_1 |1, 2, 3\rangle + a_2 |1, 3, 2\rangle + a_3 |2, 1, 3\rangle + a_4 |3, 2, 1\rangle + a_5 |2, 3, 1\rangle + a_6 |3, 1, 2\rangle.$$

(Here,  $a_i$  are complex numbers.)

△

**Exercise 2.17.** This exercise is about the number operator and states in Fock space.

- Let the number operator  $\hat{N} = \sum_P c_P^\dagger c_P$  be given. Let  $|\Phi_k\rangle = |123 \dots k\rangle$  (a Slater determinant). Compute  $\langle \Phi_k | \hat{N} | \Phi_k \rangle$ .
- Let  $\alpha > 1$  be given, and define

$$|\Psi\rangle = C \sum_k \frac{1}{k^{\alpha+2}} |\Phi_k\rangle,$$

where  $C$  is a normalization constant. Compute this constant. Is  $|\Psi\rangle$  an eigenfunction of  $\hat{N}$ ? What can you say about the number of particles in the state  $|\Psi\rangle$ ?

- Compute the expectation value of the number operator. Discuss as function of  $\alpha$ . (Hint: Riemann Zeta function.)

△

**Exercise 2.18.** In this exercise, we dig deeper into creation and annihilation operators, defining a creation operator  $c^\dagger(\psi)$  for any single-particle function  $\psi$ .

Let  $\{\phi_P\}$  be a SPF basis, with corresponding creation and annihilation operators  $\{c_P^\dagger\}$  and  $\{c_P\}$ , respectively.

Let  $\psi$  be a single-particle function. We define  $c^\dagger(\psi)$  as the operator

$$c^\dagger(\psi) \equiv \sum_P \langle \phi_P | \psi \rangle c_P^\dagger. \quad (2.105)$$

- a) Show that  $c^\dagger(\phi_P) = c_P^\dagger$ .
- b) Show that  $c^\dagger(\psi) |-\rangle = |\psi\rangle$ .
- c) For a Slater determinant  $|\phi_{P_1} \cdots \phi_{P_N}\rangle$ , show that

$$c^\dagger(\psi) |\phi_{P_1} \cdots \phi_{P_N}\rangle = |\psi \phi_{P_1} \cdots \phi_{P_N}\rangle.$$

- d) Given  $N$  arbitrary functions  $\psi_I$ , show that

$$c^\dagger(\psi_1) c^\dagger(\psi_2) \cdots c^\dagger(\psi_N) |-\rangle = |\psi_1 \psi_2 \cdots \psi_N\rangle.$$

- e) Show that  $\{c(\psi), c^\dagger(\psi')\} = \langle \psi | \psi' \rangle$ .
- f) The definition of  $c^\dagger(\psi)$  as stated in Eq. (2.105) is in terms of a given orthonormal basis. Potentially, the definition may depend on this choice of basis. Show that the definition is in fact independent of this basis, i.e., if  $\{\chi_P\}$  is another orthonormal basis with creation operators  $d_P^\dagger$ , then

$$c^\dagger(\psi) = \sum_P \langle \chi_P | \psi \rangle d_P^\dagger. \quad (2.106)$$

Hint: express  $d_P^\dagger$  in terms of the  $c_P^\dagger$ .

*Remark:* An alternative definition that is independent of basis is, for any  $|\Psi\rangle \in L_N^2$ ,

$$c^\dagger(\psi) |\Psi\rangle \equiv \sqrt{(N+1)!} \mathcal{A}_{N+1} [|\psi\rangle \otimes |\Psi\rangle], \quad (2.107)$$

where  $\mathcal{A}_{N+1}$  is the antisymmetrization operator on  $L_{N+1}^2$ .

△

**Exercise 2.19.** In this exercise we introduce field creation and annihilation operators. The *field creation operator*  $\psi^\dagger(x)$  creates a delta function centered at the point  $x$  in configuration space, i.e.,

$$\psi^\dagger(x) |-\rangle = |x\rangle, \quad \langle x' | x \rangle = \delta(x' - x). \quad (2.108)$$

For example, for electronic systems,  $\delta(x - x') \equiv \delta^3(\vec{r} - \vec{r}') \delta_{\sigma - \sigma', 0}$ , where  $\delta^3$  is the three-dimensional Dirac delta.

In this exercise the following relation may be useful:

$$\delta(x - x') = \sum_P \phi_P(x) \phi_P(x') = \sum_P \langle x | \phi_P \rangle \langle \phi_P | x' \rangle. \quad (2.109)$$

Here  $\{\phi_P\}$  is any orthonormal single-particle basis.

In terms of the definition of a creation operator for an arbitrary single-particle function given in Exercise 2.18, we have the following *definition*:

$$\psi^\dagger(x) \equiv \sum_P \langle \phi_P | x \rangle c_P^\dagger = \sum_P \phi_P(x)^* c_P^\dagger. \quad (2.110)$$

Taking the Hermitian adjoint we get,

$$\psi(x) = \sum_P \phi_P(x) c_P. \quad (2.111)$$

Now to the exercises:

a) Show that

$$\{\psi(x), \psi^\dagger(x')\} = \delta(x - x'), \quad (2.112)$$

$$\{\psi(x), \psi(x')\} = 0. \quad (2.113)$$

b) Show that

$$c^\dagger(f) = \int f(x) \psi^\dagger(x) dx, \quad (2.114)$$

$$c(f) = \int f(x)^* \psi(x) dx, \quad (2.115)$$

c) Define an operator  $\hat{\rho}(x) = \psi^\dagger(x) \psi(x)$ , depending parametrically on  $x$ . Let  $|\Phi\rangle = |\phi_1 \cdots \phi_N\rangle$  be a Slater determinant (with orthonormal single-particle functions). Compute

$$\rho(x) := \langle \Phi | \hat{\rho}(x) | \Phi \rangle, \quad (2.116)$$

showing that

$$\rho(x) = \sum_I |\phi_I(x)|^2. \quad (2.117)$$

c) (You need to know the representation of one-body operators in second quantization to do this part.) Let  $v = v(x)$  be a one-body potential, and let  $\hat{V} = \sum_i \hat{v}(i)$  be the corresponding manybody operator. Show that

$$\hat{V} = \int v(x) \hat{\rho}(x) dx. \quad (2.118)$$

Show that

$$\langle \Psi | \hat{V} | \Psi \rangle = \int v(x) \rho(x) dx. \quad (2.119)$$

△

## 2.3 Representation of operators

### 2.3.1 What we will prove

In this section, we shall demonstrate the following representation of one-body operators:

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i) = \sum_{pq} \langle p | \hat{h} | q \rangle c_p^\dagger c_q. \quad (2.120)$$

Note that the last expression *does not contain*  $N$  explicitly. Here,  $|p\rangle$  is a single-particle function – it is the “Slater determinant”  $\phi_p(x)$ . The number  $\langle p | \hat{h} | q \rangle$  is the matrix element of the single-particle operator  $\hat{h}$  in the given one-particle basis,

$$\langle p | \hat{h} | q \rangle = \int dx \phi_p(x)^* \hat{h} \phi_q(x). \quad (2.121)$$

Eq. (??) gives a nice image of how  $\hat{H}_0$  acts on a basis function: each term in the sum manipulates the Slater determinant’s occupied orbitals and weighs it with a matrix element. Simple, and not at all obvious from the “single quantized form”.

We shall also prove the following formula for the two-body operator:

$$\hat{W} = \sum_{(i,j)}^N \hat{w}(i,j) = \frac{1}{2} \sum_{pqrs} w_{rs}^{pq} c_p^\dagger c_q^\dagger c_s c_r, \quad (2.122)$$

where the ordering of the annihilation operators should be noted. Here,

$$w_{rs}^{pq} = \int dx_1 \int dx_2 \phi_p(x_1)^* \phi_q(x_2)^* w(x_1, x_2) \phi_r(x_1) \phi_s(x_2) \quad (2.123)$$

is a matrix element using tensor product two-body functions, *not* Slater determinants. Using Slater determinant matrix elements we in fact have a similar expansion,

$$\hat{W} = \frac{1}{4} \sum_{pqrs} \langle pq | \hat{w} | rs \rangle c_p^\dagger c_q^\dagger c_s c_r, \quad (2.124)$$

where thus the matrix elements *are antisymmetric*, computed as a matrix element using two-body Slater determinants. It is a fact that

$$\langle pq | \hat{w} | rs \rangle = w_{rs}^{pq} - w_{sr}^{pq}. \quad (2.125)$$

A word of warning: notation for two-body matrix elements is notoriously varying between sources. Some authors use the notation  $\langle \phi_p \phi_q | \hat{w} | \phi_r \phi_s \rangle$  for the matrix element  $w_{rs}^{pq}$ , which is *not* antisymmetric. In our case, the notation clashes with the Slater determinant matrix element, but we will still sin in this respect. Some authors write  $\langle \phi_p \phi_q | \hat{w} | \phi_r \phi_s \rangle_{AS}$  for the anti-symmetric Slater-determinant matrix element (and sometimes we will too).

There are many different conventions in use. A general piece of advice is to make sure in a given context what are the conventions.

The proofs given in this section borrow heavily from [5].

A very useful result in general, is the following lemma:

**Lemma 2.1.** *The action of the operator  $c_q^\dagger c_p$  on a Slater determinant (built from the orthonormal  $\{\phi_r\}$ ) is as follows: it replaces  $\phi_p$  with  $\phi_q$  in the Slater determinant, with no sign change, or gives zero if  $\phi_p$  is not occupied.*

Similarly,  $c_{q_1}^\dagger c_{q_2}^\dagger c_{p_2} c_{p_1}$  replaces  $p_1$  with  $q_1$ , and  $p_2$  with  $q_2$ , or gives zero if one of  $p_1$  or  $p_2$  is not present in the determinant.

*Proof.* See Exercise 2.20. □



### 2.3.2 One-body operators

We prove Eq. (2.120) by showing that the actions of the left- and right-hand sides on an arbitrary Slater determinant agree. Let therefore  $\{\phi_\mu\}$  be a single-particle basis as usual.

Consider the action of  $\hat{H}_0 = \sum_i \hat{h}(i)$  on an arbitrary Slater determinant:

$$\begin{aligned}\hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \frac{1}{\sqrt{N!}} \left( \sum_i \hat{h}(i) \right) \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_{\sigma} \phi_{\nu_1}(x_1) \cdots \phi_{\nu_N}(x_N) \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_{\sigma} \left( \sum_i \hat{h}(i) \right) \phi_{\nu_1}(x_1) \cdots \phi_{\nu_N}(x_N) \\ &= |(\hat{h}\phi_{\nu_1}), \phi_{\nu_2}, \dots, \phi_{\nu_N}\rangle + |\phi_{\nu_1}, (\hat{h}\phi_{\nu_2}), \dots, \phi_{\nu_N}\rangle + \cdots + |\phi_{\nu_1}, \phi_{\nu_2}, \dots, (\hat{h}\phi_{\nu_N})\rangle\end{aligned}\quad (2.126)$$

Here, we used that  $\hat{P}_{\sigma}$  commutes with  $\hat{H}_0$ .

Consider the operator  $\hat{h}$  acting on a single-particle function  $\phi_\mu$ . The result,  $\psi$ , can be expanded in the basis:

$$\psi(x) = \hat{h}\phi_\mu(x) = \sum_{\nu} \phi_{\nu}(x) \langle \nu | \hat{h} | \mu \rangle. \quad (2.127)$$

We insert this expansion:

$$\begin{aligned}\hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= |(\hat{h}\phi_{\nu_1}), \phi_{\nu_2}, \dots, \phi_{\nu_N}\rangle + |\phi_{\nu_1}, (\hat{h}\phi_{\nu_2}), \dots, \phi_{\nu_N}\rangle + \cdots + |\phi_{\nu_1}, \phi_{\nu_2}, \dots, (\hat{h}\phi_{\nu_N})\rangle \\ &= \sum_{\nu} \langle \nu | \hat{h} | \mu_1 \rangle |\nu \mu_2, \dots, \mu_N\rangle + \sum_{\nu} \langle \nu | \hat{h} | \mu_2 \rangle |\mu_1 \nu \mu_3, \dots, \mu_N\rangle + \cdots + \sum_{\nu} \langle \nu | \hat{h} | \mu_N \rangle |\mu_1 \mu_2, \dots, \nu\rangle\end{aligned}\quad (2.128)$$

Now, we note that

$$|\mu_1, \dots, \mu_{j-1} \nu \mu_{j+1}, \dots, \mu_N\rangle = c_{\nu}^{\dagger} c_{\mu_j} |\mu_1, \dots, \mu_N\rangle, \quad (2.129)$$

which we plug in:

$$\begin{aligned}\hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \sum_{\nu} \langle \nu | \hat{h} | \mu_1 \rangle |\nu \mu_2, \dots, \mu_N\rangle + \sum_{\nu} \langle \nu | \hat{h} | \mu_2 \rangle |\mu_1 \nu \mu_3, \dots, \mu_N\rangle + \cdots + \sum_{\nu} \langle \nu | \hat{h} | \mu_N \rangle |\mu_1 \mu_2, \dots, \nu\rangle \\ &= \left[ \sum_{\nu} \langle \nu | \hat{h} | \mu_1 \rangle c_{\nu}^{\dagger} c_{\mu_1} + \sum_{\nu} \langle \nu | \hat{h} | \mu_2 \rangle c_{\nu}^{\dagger} c_{\mu_2} + \cdots + \sum_{\nu} \langle \nu | \hat{h} | \mu_N \rangle c_{\nu}^{\dagger} c_{\mu_N} \right] |\mu_1, \dots, \mu_N\rangle.\end{aligned}\quad (2.130)$$

Finally, we note that  $c_{\mu} |\mu_1 \cdots \mu_N\rangle = 0$  whenever  $\mu \notin \vec{\mu}$ , so we may extend the summation over  $\mu_j$  to all of  $\mu$ , resulting in:

$$\begin{aligned}\hat{H}_0 |\phi_{\mu_1}, \phi_{\mu_2}, \dots, \phi_{\mu_N}\rangle &= \left[ \sum_{\nu} \langle \nu | \hat{h} | \mu_1 \rangle c_{\nu}^{\dagger} c_{\mu_1} + \sum_{\nu} \langle \nu | \hat{h} | \mu_2 \rangle c_{\nu}^{\dagger} c_{\mu_2} + \cdots + \sum_{\nu} \langle \nu | \hat{h} | \mu_N \rangle c_{\nu}^{\dagger} c_{\mu_N} \right] |\mu_1, \dots, \mu_N\rangle \\ &= \sum_{\mu\nu} \langle \nu | \hat{h} | \mu \rangle c_{\nu}^{\dagger} c_{\mu} |\mu_1, \dots, \mu_N\rangle.\end{aligned}\quad (2.131)$$

Since  $|\mu_1, \dots, \mu_N\rangle$  was an arbitrary Slater determinant, we have proven Eq. (2.120).

### 2.3.3 Two-body operators

The operator  $\hat{W} = \sum_{i < j} \hat{w}(i, j)$  is a two-body operator. The operator  $\hat{w}(1, 2)$  is thus an operator on  $L^2(X^2)$  that is completely characterized by its action on a basis: the tensor products  $\phi_{\mu_1}(x_1)\phi_{\mu_2}(x_2)$ . Thus,

$$\hat{w}(1, 2)\phi_{\mu_1}(x_1)\phi_{\mu_2}(x_2) = \sum_{\nu_1 \nu_2} w_{\mu_1 \mu_2}^{\nu_1 \nu_2} \phi_{\nu_1}(x_1)\phi_{\nu_2}(x_2), \quad (2.132)$$

where the matrix elements  $w_{\mu_1 \mu_2}^{\nu_1 \nu_2}$  are given by the formula (2.123). There is nothing special about the indices  $(1, 2)$ , it may just as well be  $(i, j)$ . Note also the symmetry property

$$w_{\mu_1 \mu_2}^{\nu_1 \nu_2} = w_{\mu_2 \mu_1}^{\nu_2 \nu_1}.$$

As for the one-body case,  $\hat{W}$  commutes with  $\hat{P}_\sigma$ , and we get, using Eq. (2.132),

$$\begin{aligned} \hat{W} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_\sigma \left[ \sum_{i < j} \hat{w}(i, j) \phi_{\mu_1}(x_1) \cdots \phi_{\mu_N}(x_N) \right] \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_\sigma \left[ \sum_{i < j} \hat{w}(i, j) \phi_{\mu_1}(x_1) \cdots \phi_{\mu_N}(x_N) \right] \\ &= \frac{1}{\sqrt{N!}} \sum_{\sigma} (-1)^{|\sigma|} \hat{P}_\sigma \left[ \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} \phi_{\mu_1} \cdots \phi_{\nu_1}(x_i) \cdots \phi_{\nu_2}(x_j) \cdots \phi_{\mu_N}(x_N) \right] \quad (2.133) \\ &= \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} |\phi_{\mu_1} \cdots \phi_{\nu_1} \cdots \phi_{\nu_2} \cdots \phi_{\mu_N}\rangle \\ &= \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle. \end{aligned}$$

Here, we used Lemma 2.1 about replacement behavior of the  $c^\dagger c^\dagger c c$  product. We are currently summing over  $\mu_i$  and  $\mu_j$ , such that  $i < j$ . Including  $i = j$  gives zero contribution (why?), and including  $j > i$  gives equal contribution:

$$\begin{aligned} \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle &= - \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_i} c_{\mu_j} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle \\ &= \sum_{i < j} \sum_{\nu_1 \nu_2} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_2}^\dagger c_{\nu_1}^\dagger c_{\mu_i} c_{\mu_j} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle = \sum_{j < i} \sum_{\nu_1 \nu_2} w_{\mu_j \mu_i}^{\nu_2 \nu_1} c_{\nu_2}^\dagger c_{\nu_1}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle \quad (2.134) \\ &= \sum_{j < i} \sum_{\nu_2 \nu_1} w_{\mu_j \mu_i}^{\nu_2 \nu_1} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle = \sum_{j < i} \sum_{\nu_2 \nu_1} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle \end{aligned}$$

Here, we used the anticommutators and symmetry of the matrix elements. Assembling the two contributions,

$$\hat{W} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle = \frac{1}{2} \sum_{ij} \sum_{\nu_2 \nu_1} w_{\mu_i \mu_j}^{\nu_1 \nu_2} c_{\nu_1}^\dagger c_{\nu_2}^\dagger c_{\mu_j} c_{\mu_i} |\phi_{\mu_1} \cdots \phi_{\mu_N}\rangle. \quad (2.135)$$

We note that the sum over  $ij$  is really a sum over two occupied orbitals  $\mu_i$  and  $\mu_j$ . We can therefore extend the sum to all unoccupied orbitals as well, since  $c_\alpha |\vec{\mu}\rangle$  gives zero contributions for such orbitals. Thus, Eq. (2.122) is proven.

We leave the proof of the antisymmetrized version as an exercise, ??.

### 2.3.4 Diagonalizing a one-body operator

The Hamiltonian is often of the form

$$\hat{H} = \hat{H}_0 + \hat{W}, \quad (2.136)$$

where  $\hat{H}_0$  is “exactly solvable”, in the sense that we can find a complete set of eigenvectors analytically. Almost always,  $\hat{H}_0 = \sum_i \hat{h}(i)$  is a one-body operator. The eigenfunctions are a natural starting point for treating the full problem  $\hat{H}$ .

One often refers to  $\hat{H}_0$  as the “zero order Hamiltonian”, with  $\hat{W}$  “the perturbation”, even if one does not use perturbation theory. One thinks of  $\hat{H}_0$  as a first approximation.

Suppose we have a complete set of eigenvectors to  $\hat{h}$ ,

$$\hat{h}\phi_Q(x) = \epsilon_Q\phi_Q(x). \quad (2.137)$$

Forming the inner product with  $\phi_P$  we obtain

$$\langle \phi_P | \hat{h} | \phi_Q \rangle = \delta_{PQ} \epsilon_P. \quad (2.138)$$

The complete set of eigenvectors can be used to form a complete set of Slater determinant basis functions for  $N$ -fermion space, and for Fock space,

$$|\phi_{P_1}\phi_{P_2}\cdots\phi_{P_N}\rangle = c_{P_1}^\dagger \cdots c_{P_N}^\dagger |-\rangle, \quad (2.139)$$

where we have introduced creation and annihilation operators relative to the basis. We then obtain the following expression for  $\hat{H}_0$ :

$$\hat{H}_0 = \sum_P \epsilon_P c_P^\dagger c_P. \quad (2.140)$$

Acting on a Slater determinant, we see that (prove this!)

$$c_{P_i}^\dagger c_{P_i} |\phi_{P_1}\cdots\phi_{P_N}\rangle = |\phi_{P_1}\cdots\phi_{P_N}\rangle. \quad (2.141)$$

and thus that

$$\hat{H}_0 |\phi_{P_1}\cdots\phi_{P_N}\rangle = \left( \sum_i \epsilon_{P_i} \right) |\phi_{P_1}\cdots\phi_{P_N}\rangle, \quad (2.142)$$

Since the Slater determinants are complete, we have found that *the Slater determinants form a complete set of eigenvectors for  $\hat{H}_0$ , with eigenvalues given by the sum of the single-particle eigenvalues.*

### 2.3.5 Exercises

**Exercise 2.20.** Prove Lemma 2.1

△

**Exercise 2.21.** This exercise tests your understanding of second-quantized operators. It also gives a simple example of a “plane-wave basis” for a toy model. The toy model can be a “quantum ring”, electrons living on a circle. (For this, think of  $\theta = 2\pi x$  as an angle.)

Recall that for a square-integrable function  $f : [0, 1] \rightarrow \mathbb{C}$ , we can find a unique Fourier series,

$$f(x) = \sum_{n \in \mathbb{Z}} f_n e^{i2\pi n x}.$$

This series always exists.

- a) Show that  $\varphi_n(x) = e^{i2\pi nx}$  forms a set of orthonormal functions over the interval  $[0, 1]$ . (Given that the expansion of  $f(x)$  always exists, we then have a basis.) Show that

$$f_n = \int_0^1 e^{-i2\pi nx} f(x) dx.$$

- b) We consider the single-particle functions  $\phi_{n,\sigma}(x, s) = \varphi_n(x)\chi_\sigma(s)$ , where  $\chi_\sigma$  is the spin basis function for the spin direction  $\sigma \in \{+1, -1\}$ .

We let  $P = (n, \sigma)$ .

Verify that  $\{\phi_P\}$  is an orthonormal set,  $\langle \phi_P | \phi_Q \rangle = \delta_{PQ}$ .

- c) Let  $\hat{t} = -\frac{1}{2} \frac{\partial^2}{\partial x^2}$  be the kinetic energy operator, and consider the  $N$ -body operator

$$\hat{T} = \sum_i \hat{t}(i).$$

Compute the matrix elements  $\langle \phi_P | \hat{T} | \phi_Q \rangle$ . Write down the second-quantized operator  $\hat{T}$  using these and the creation and annihilation operators  $c_{n\sigma}^{(\dagger)}$ .

- d) Let  $\hat{v} = \cos(2\pi mx)$  be a potential, where  $m$  is an integer. Compute the matrix elements  $\langle \phi_P | \hat{v} | \phi_Q \rangle$ . Write down the second-quantized operator corresponding to  $\hat{V} = \sum_i \hat{v}(i)$  using these and the creation and annihilation operators  $c_{n\sigma}^{(\dagger)}$ .

- d) Consider now an arbitrary periodic potential  $v(x)$ . Compute the second-quantized operator  $\hat{V} = \sum_i \hat{v}(i)$  in terms of the Fourier coefficients of  $v(x)$ .

- e) We now consider a two-body operator. Let  $w(x)$  be a periodic function over  $x \in [0, 1]$ , and let

$$\hat{w}(1, 2) = w(x_1 - x_2), \quad \hat{W} = \sum_{i < j=1}^N \hat{w}(i, j).$$

Compute the (non-antisymmetrized) matrix elements  $\langle \phi_{n_1\sigma_1} \phi_{n_2\sigma_2} | \hat{w} | \phi_{n_3\sigma_3} \phi_{n_4\sigma_4} \rangle$  in terms of the Fourier coefficients of  $w$ . Compute the second-quantized form of  $\hat{W}$ . Make sure your result sums over as few spins and as few Fourier mode indices as possible, i.e., collapse any Kronecker deltas you encounter.

△

**Exercise 2.22.** It is recommended to do Exercise 2.18 before this exercise.

Let  $U = [U_{QP}]$  be a unitary matrix. There exists an  $S = -S^H$  such that  $U = e^S$ . (You do not have to show this.) Let  $U$  define a unitary change of single-particle basis,

$$\psi_P = \sum_Q \phi_Q U_{QP}. \quad (2.143)$$

Let  $c_P^\dagger = c^\dagger(\phi_P)$  and  $d_P^\dagger = c^\dagger(\psi_P)$ .

- a) Define  $\hat{S} = \sum_{PQ} S_{PQ} c_P^\dagger c_Q$ . Show that  $\hat{U} \equiv e^{\hat{S}}$  is a unitary operator. Is  $\hat{U}$  a one-body operator? If not, how many particles does it act on simultaneously?

b) Compute  $[\hat{S}, c_P^\dagger]$ ,  $[\hat{S}, [\hat{S}, c_P^\dagger]]$ , and by iterating, the  $n$ -fold nested commutator  $[\hat{S}, [\dots [\hat{S}, c_P^\dagger] \dots]]$ .

c) Show that

$$d_P^\dagger = e^{\hat{S}} c_P^\dagger e^{-\hat{S}}.$$

Hint: recall the Baker–Campbell–Hausdorff formula

$$e^A B e^{-A} = B + [A, B] + \frac{1}{2}[A, [A, B]] + \dots + \frac{1}{n!}[A, [A, \dots [A, B] \dots]] + \dots \quad (2.144)$$

d) Show that

$$e^{\hat{S}} |\phi_{P_1} \dots \phi_{P_N}\rangle = |\psi_{P_1} \dots \psi_{P_N}\rangle, \quad (2.145)$$

i.e., that  $e^{\hat{S}}$  is the manybody basis change operator corresponding to the basis change in Eq. (2.143).

△

**Exercise 2.23.** It may be useful to do Exercise 2.22 before this exercise.

In this exercise, we use the notation  $U = [U_{QP}]$  to denote a matrix with indices that are both occupied and virtual. If we specify, say  $V = [V_{IA}]$ , the row index is occupied, and the column index virtual.

Here, we prove *Thouless Theorem*: Let  $|\Phi\rangle$  be a Slater determinant,

$$|\Phi\rangle = |\phi_1 \phi_2 \dots \phi_N\rangle, \quad (2.146)$$

with orthonormal  $\{\phi_P\}$ . Then any other Slater determinant  $|\Psi\rangle$  that satisfies  $\langle\Phi|\Psi\rangle = 1$  can be written

$$|\Psi\rangle = e^{\hat{T}} |\Phi\rangle, \quad \hat{T} = \sum_{IA} t_{IA} c_A^\dagger c_I, \quad (2.147)$$

where  $\hat{T}$  is called a *singles excitation operator* (see coupled-cluster theory). Conversely,  $|\Psi\rangle = e^{\hat{T}} |\Phi\rangle$  is always a Slater determinant with  $\langle\Phi|\Psi\rangle = 1$ .

a) Let  $T = [t_{PQ}]$  be the matrix, whose only non-zero elements are  $t_{IA}$ . That is,  $t_{AI} = t_{AB} = t_{IJ} = 0$ . Show that  $e^{\hat{T}} = 1 + \hat{T}$ , and that  $e^{-\hat{T}} = 1 - \hat{T}$ .

Let  $\psi_I = \phi_I + \sum_A t_{IA} \phi_A$ , with creation operators  $d_I^\dagger$ . (In general,  $\langle\psi_I|\psi_J\rangle \neq \delta_{IJ}$ .) Show that  $d_I^\dagger = e^{\hat{T}} c_I^\dagger e^{-\hat{T}}$ .

b) Show that  $|\Psi\rangle = e^{\hat{T}} |\Phi\rangle$  is a Slater determinant and that  $\langle\Phi|\Psi\rangle = 1$ .

c) Let  $|\Psi\rangle$  be a Slater determinant with  $\langle\Phi|\Psi\rangle = 1$ . Assume that the single-particle functions  $\psi_I$  in  $|\Psi\rangle$  are orthonormal, and explain that there exists a matrix  $U = [U_{QI}]$ , with  $U^H U = 1$ , such that

$$\psi_I = \sum_Q \phi_Q U_{QI}.$$

Show that

$$\langle\Phi|\Psi\rangle = \det([U_{IJ}]).$$

d) Show that if  $A = [A_{IJ}]$  is a matrix with  $\det(A) = 1$ , and if  $\tilde{\psi}_I = \sum_J \psi_J A_{JI}$ , then

$$|\psi_1 \cdots \psi_N\rangle = |\tilde{\psi}_1 \cdots \tilde{\psi}_N\rangle.$$

Find  $A$  such that

$$\tilde{\psi}_I = \phi_I + \sum_A t_{IA} \phi_A$$

for some  $t = [t_{IA}]$ . Explain that Thouless' Theorem is now proven.

△

**Exercise 2.24.** Prove Eq. (2.124). Start with showing Eq. (??).

△

**Exercise 2.25.** a) Let  $\hat{F} = \sum_{i=1}^N \hat{f}(i)$  be a first-quantization operator. Write down the second-quantized form of this operator. Let  $\hat{G} = \sum_{i < j} \hat{g}(i, j)$  be a general two-body operator, where  $\hat{g}(1, 2) = \hat{g}(2, 1)$ . Write down the second-quantized form.

b) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 | \hat{F} | \mu_1 \mu_2 \rangle$$

c) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 \mu_3 | \hat{F} | \mu_1 \mu_2 \mu_3 \rangle$$

d) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 | \hat{G} | \mu_1 \mu_2 \rangle$$

e) Using the fundamental anticommutator relations, compute the matrix element

$$\langle \mu_1 \mu_2 \mu_3 | \hat{G} | \mu_1 \mu_2 \mu_3 \rangle$$

f) Compute the matrix element

$$\langle \mu_1, \mu_2, \cdots, \mu_N | \hat{F} | \mu_1, \mu_2, \cdots, \mu_N \rangle$$

g) Compute the matrix element

$$\langle \mu_1, \mu_2, \cdots, \mu_N | \hat{G} | \mu_1, \mu_2, \cdots, \mu_N \rangle$$

△

**Exercise 2.26.** (Tedious, but very instructive.) In this exercise, we prove the so-called *Slater–Condon rules*: the explicit expressions of matrix elements of one- and two-body operators in a Slater determinant basis.

We do not assume any particular ordering of the occupied single-particle functions considered.

If you solved Exercise 2.25, you solved parts of this exercise.

a) Using the fundamental anticommutator relations, compute  $\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle$  and  $\langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle$  and prove that

$$\langle \vec{\mu} | \hat{H}_0 | \vec{\mu} \rangle = \sum_{i=1}^N h_{\mu_i}^{\mu_i}, \quad (2.148)$$

$$\langle \vec{\mu} | \hat{W} | \vec{\mu} \rangle = \sum_{i < j}^N \langle \mu_i \mu_j | \hat{w} | \mu_i \mu_j \rangle_{\text{AS}} = \frac{1}{2} \sum_{ij} \langle \mu_i \mu_j | \hat{w} | \mu_i \mu_j \rangle_{\text{AS}}. \quad (2.149)$$

b) Let  $\vec{v}$  be equal to  $\vec{\mu}$ , except for one occupied orbital, i.e.,

$$|\vec{v}\rangle = c_{\nu_j}^\dagger c_{\mu_j} |\vec{\mu}\rangle, \quad \nu_j \neq \mu_j. \quad (2.150)$$

Using the fundamental anticommutator relations, compute  $\langle \vec{\mu} | \hat{H}_0 | \vec{v} \rangle$  and  $\langle \vec{\mu} | \hat{W} | \vec{v} \rangle$ , and find

$$\langle \vec{\mu} | \hat{H}_0 | \vec{v} \rangle = h_{\nu_j}^{\mu_j}, \quad (2.151)$$

$$\langle \vec{\mu} | \hat{W} | \vec{v} \rangle = \sum_i \langle \mu_i \mu_j | \hat{w} | \mu_i \nu_j \rangle_{\text{AS}}. \quad (2.152)$$

c) Let  $\vec{v}$  be equal to  $\vec{\mu}$ , except for two indices, i.e.,

$$|\vec{v}\rangle = c_{\nu_k}^\dagger c_{\nu_j}^\dagger c_{\mu_j} c_{\mu_k} |\vec{\mu}\rangle, \quad j \neq k. \quad (2.153)$$

Using the fundamental anticommutator relations, compute  $\langle \vec{\mu} | \hat{W} | \vec{v} \rangle$  and find

$$\langle \vec{\mu} | \hat{H}_0 | \vec{v} \rangle = 0, \quad (2.154)$$

$$\langle \vec{\mu} | \hat{W} | \vec{v} \rangle = \langle \mu_j \mu_k | \hat{w} | \nu_j \nu_k \rangle_{\text{AS}}. \quad (2.155)$$

d) Explain that if  $\vec{v}$  differs from  $\vec{\mu}$  in *more* than two occupied functions, then  $\langle \vec{\mu} | \hat{W} | \vec{v} \rangle = 0$ .

△

**Exercise 2.27.** (This exercise adapted from an exercise by Morten Hjorth-Jensen.)

We will now consider a simple three-level problem, depicted in Figure 2.1. The single-particle states are labeled by the quantum number  $p$  and can accommodate up to two single particles, viz., every single-particle state is doubly degenerate (you could think of this as one state having spin up and the other spin down). We let the spacing between the doubly degenerate single-particle states be constant, with value  $d$ . The first state has energy  $d$ . There are only three available single-particle states,  $p = 1$ ,  $p = 2$  and  $p = 3$ , as illustrated in the figure.

a) How many two-particle Slater determinants can we construct in this space?

b) We limit ourselves to a system with only the two lowest single-particle orbits and two particles,  $p = 1$  and  $p = 2$ . We assume that we can write the Hamiltonian as

$$\hat{H} = \hat{H}_0 + \hat{H}_I,$$

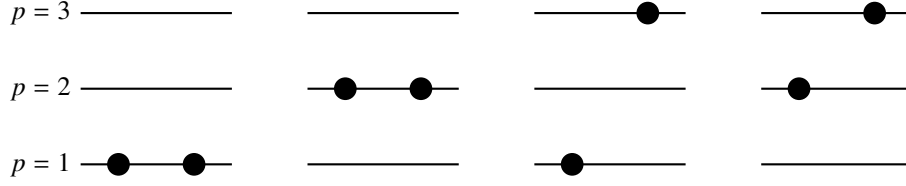


Figure 2.1: Schematic plot of the possible single-particle levels with double degeneracy. The filled circles indicate occupied particle states. The spacing between each level  $p$  is constant in this picture. We show some possible two-particle states.

and that the onebody part of the Hamiltonian with single-particle operator  $\hat{h}_0$  has the property

$$\hat{h}_0 \psi_{p\sigma} = p \times d \psi_{p\sigma},$$

where we have added a spin quantum number  $\sigma$ . We assume also that the only two-particle states that can exist are those where two particles are in the same state  $p$ , as shown by the two possibilities to the left in the figure. The two-particle matrix elements of  $\hat{H}_I$  have all a constant value,  $-g$ . Show then that the Hamiltonian matrix can be written as

$$\begin{pmatrix} 2d - g & -g \\ -g & 4d - g \end{pmatrix},$$

and find the eigenvalues and eigenvectors. What is mixing of the state with two particles in  $p = 2$  to the wave function with two-particles in  $p = 1$ ? Discuss your results in terms of a linear combination of Slater determinants.

- c) Add the possibility that the two particles can be in the state with  $p = 3$  as well and find the Hamiltonian matrix, the eigenvalues and the eigenvectors. We still insist that we only have two-particle states composed of two particles being in the same level  $p$ . You can diagonalize numerically your  $3 \times 3$  matrix.

This simple model catches several birds with a stone. It demonstrates how we can build linear combinations of Slater determinants and interpret these as different admixtures to a given state. It represents also the way we are going to interpret these contributions. The two-particle states above  $p = 1$  will be interpreted as excitations from the ground state configuration,  $p = 1$  here. The reliability of this ansatz for the ground state, with two particles in  $p = 1$ , depends on the strength of the interaction  $g$  and the single-particle spacing  $d$ . Finally, this model is a simple schematic ansatz for studies of pairing correlations and thereby superfluidity/superconductivity in fermionic systems.

△



## 2.4 Many-fermion systems: examples

In this section, we consider some typical many-fermion systems to which our theory can be applied.

### 2.4.1 Electronic systems

The book by Stanley Raimes is a good supporting book here.

Electrons move about in  $\mathbb{R}^3$ . For some models it's useful to consider  $\mathbb{R}^d$ , with  $d = 1, 2$  as well. Additionally, they have electron spin, i.e., they are spin- $\frac{1}{2}$  particles. Thus, the coordinates of a single electron are

$$x = (\vec{r}, s), \quad \vec{r} \in \mathbb{R}^d, \quad s \in \{\uparrow, \downarrow\}.$$

The discrete spin variable can have a number of arbitrary denominations, e.g.,  $s \in \{1, -1\}$ . Chemists often talk about  $\alpha$  and  $\beta$  spin.

Thus, a single-particle function  $\psi(x)$  has two components, one for each value of the spin:

$$\psi(x) = \psi(\vec{r}, s) = \psi_{\uparrow}(\vec{r})\chi_{\uparrow}(s) + \psi_{\downarrow}(\vec{r})\chi_{\downarrow}(s).$$

Here,  $\chi_{\sigma}$  is a spinor basis function, and often taken to be the eigenvectors of  $\sigma_z$ , the  $z$ -component of the Pauli matrices,

$$\vec{\sigma} = \begin{bmatrix} \sigma_x \\ \sigma_y \\ \sigma_z \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \\ \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \end{bmatrix} \quad (2.156)$$

i.e.,

$$\chi_{\sigma}(s) = \delta_{\sigma,s}.$$

Let us consider a single-particle basis  $\{\phi_{\mu}(x)\}$ . We here use a greek index to label the state. In many cases we assume that  $\phi_{\mu}$  is an eigenfunction of the spin projection along the  $z$ -axis, i.e.,  $\phi_{\mu}(x) = \varphi_p(\vec{r})\chi_{\sigma}(s)$ , so that  $\mu = (p, \sigma)$ . A Slater determinant built from *spin-orbitals* can thus be pictured like this: [picture here]

Assuming that relativistic effects can be ignored, the Hamiltonian of the many-electron system is on the form

$$\hat{H} = \hat{T} + \hat{W} + \hat{V}, \quad (2.157)$$

where the kinetic energy is

$$\hat{T} = \sum_{i=1}^N \hat{t}(i), \quad \hat{t} = -\frac{\hbar^2}{2m_e} \nabla^2. \quad (2.158)$$

The inter-electronic Coulomb force is

$$\hat{W} = \frac{1}{2} \sum_{i \neq j}^N \hat{w}(i, j), \quad w(\vec{r}_1, \vec{r}_2) = \frac{e^2}{|\vec{r}_1 - \vec{r}_2|} \quad (2.159)$$

The term  $\hat{V}$  represents external forces acting on the system. For a molecule in the Born–Oppenheimer approximation, we have

$$\hat{V} = \sum_{i=1}^N v(i), \quad v(\vec{r}) = \sum_{\alpha} \frac{-e^2 Z_{\alpha}}{|\vec{r} - \vec{R}_{\alpha}|}, \quad (2.160)$$

where the nuclei are indexed using  $\alpha$ , and have positions  $\vec{R}_\alpha$  and charges  $eZ_\alpha$ .

### 2.4.2 The $N$ -electron parabolic quantum dot in one dimension

In this section, we consider an example, applying our current level of knowledge to a non-trivial model system, namely that of a system of  $N$  interacting electrons in a Harmonic oscillator trap. For simplicity, we consider a one-dimensional model.

Single-particle space: in one spatial dimension, a single-particle wavefunction is a function

$$\psi : \mathbb{R} \times \{+1, -1\} \rightarrow \mathbb{C}, \quad (2.161)$$

i.e.,  $\psi(x) = \psi(z, s)$ , where  $z \in \mathbb{R}$  is the space coordinate and  $s = \pm$  is the spin orientation along some axis.

The Hamiltonian has the form

$$\hat{H} = \hat{H}_0 + \hat{W}, \quad (2.162)$$

with, after a suitable scaling of units,

$$\hat{H}_0 = \sum_{i=1}^N \hat{h}(i), \quad \hat{h} = -\frac{1}{2} \frac{\partial^2}{\partial z^2} + \frac{1}{2} \omega z^2. \quad (2.163)$$

Note that the single-particle operator  $\hat{h}$  does not depend on spin. This is just the Harmonic oscillator, with eigenfunctions  $\varphi_n(z)$  given by Hermite polynomials multiplied by a Gaussian, giving the single-particle basis functions

$$\phi_P(x) = \varphi_n(z) \chi_\sigma(s), \quad (2.164)$$

where  $\chi_s$  are the spinor basis functions,  $s = \pm$ . Thus,  $P = (n, s)$  is the single-particle index. The eigenvalues are

$$\epsilon_n = \omega(n + \frac{1}{2}). \quad (2.165)$$

We introduce creation and annihilation operator for the basis functions, and a complete set of Slater determinant basis functions for Fock space. We obtain for  $\hat{H}_0$ ,

$$\hat{H}_0 = \sum_n \epsilon_n \sum_\sigma c_{n,\sigma}^\dagger c_{n,\sigma}. \quad (2.166)$$

The eigenvalues of  $\hat{h}$  are evenly spaced, but doubly degenerate. A diagram like Fig. 2.2 this is therefore handy for illustrating the system and the Slater determinants constructed from the single-particle basis.

Let us consider the interaction,

$$\hat{W} = \sum_{i < j}^N \hat{w}(i, j). \quad (2.167)$$

Here,  $w(x_1, x_2) = w(\vec{r}_1, \vec{r}_2)$  is the interaction potential between the electrons. In 3D this is the Coulomb interaction, while in 1D and 2D models, which are only *effective* models, and we introduce a screening parameter  $\delta$  to remove the Coulomb singularity, but retain the asymptotic behavior at large distances, i.e.,

$$w(z_1, z_2) = \frac{\lambda}{\sqrt{(z_1 - z_2)^2 + \delta^2}}. \quad (2.168)$$

Here,  $\lambda$  is a dimensionless constant arising from the change to dimensionless units.

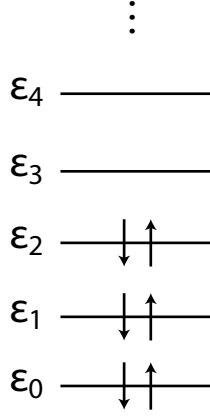


Figure 2.2: Figure of one-particle energy levels. A 6-electron Slater determinant is illustrated.

The operator  $\hat{W}$  has the general form

$$\hat{W} = \frac{1}{2} \sum_{PQRS} \langle \phi_P \psi_Q | \hat{w} | \phi_R \phi_S \rangle c_P^\dagger c_Q^\dagger c_S c_R, \quad (2.169)$$

where we need to evaluate the brackets. We introduce the quantum numbers  $P = (n_P, \sigma_P)$  etc, and compute

$$\begin{aligned} \langle \phi_P \psi_Q | \hat{w} | \phi_R \phi_S \rangle &= \int dx_1 \int dx_2 \phi_P(x_1) \phi_Q(x_2) w(x_1, x_2) \phi_R(x_1) \phi_S(x_2) \\ &= \sum_{s_1 s_2} \chi_{\sigma_P}(s_1) \chi_{\sigma_Q}(s_2) \chi_{\sigma_R}(s_1) \chi_{\sigma_S}(s_2) \int dz_1 \int dz_2 \varphi_{n_P}(z_1) \varphi_{n_Q}(z_2) w(z_1, z_2) \varphi_{n_R}(z_1) \varphi_{n_S}(z_2) \\ &\equiv \delta_{\sigma_P \sigma_R} \delta_{\sigma_Q \sigma_S} (n_P n_Q | n_R n_S). \end{aligned} \quad (2.170)$$

We note here that  $\sigma_P = \sigma_R$  and  $\sigma_Q = \sigma_S$  is required for a non-vanishing bracket. We have also defined the *spatial integrals*  $(n_1 n_2 | n_3 n_4)$ . Plugging back into  $\hat{W}$ , we obtain

$$\hat{W} = \frac{1}{2} \sum_{n_1 n_2 n_3 n_4} (n_1 n_2 | n_3 n_4) \sum_{\sigma \tau} c_{n_1 \sigma}^\dagger c_{n_2 \tau}^\dagger c_{n_4 \tau} c_{n_3 \sigma}. \quad (2.171)$$

### 2.4.3 The nuclear manybody problem

We will briefly mention the structure of the nuclear manybody problem in a nonrelativistic approximation. For an exposition, see for example the textbook by Bernard L. Cohen, “Concepts of nuclear physics”.

The atomic nucleus consists of nucleons, i.e., protons and neutrons. Each have three quarks: the proton has 2 up quarks and 1 down quark, and the neutron has 1 up quark and 2 down quarks. They are thus composite particles described by QCD. Yet, a nonrelativistic model Hamiltonian is still very useful.

Neutrons and protons are spin- $\frac{1}{2}$  particles, and are modeled as the *same* species differing only in the value of the projection of *isobaric spin*. This is a feature of the nuclear shell model, where the Pauli exclusion principle is prominent. When isobaric spin is  $\tau = \uparrow$ , the nucleon is a proton, and when  $\tau = \downarrow$  it is a neutron. Thus, the single-particle coordinates are

$$x = (\vec{r}, s, t).$$

The Hamiltonian is basically *unknown* and can be written as

$$\begin{aligned}\hat{H} &= \hat{T} + \hat{V}_{\text{NN}} + \hat{V}_{\text{NNN}} + \cdots \\ &= \sum_{i=1}^A \hat{t}(i) + \sum_{i<j}^A v_{\text{NN}}(i, j) + \sum_{i<j<k}^A v_{\text{NN}}(i, j, k) + \cdots\end{aligned}\quad (2.172)$$

where  $\hat{V}_{\text{NN}}$  is a two-body potential, etc. It is typical today to include up to three-body forces in the description.

The total number of nucleons is denoted  $A = N + Z$ , where  $N$  is the number of neutrons. (The Hamiltonian preserves  $N$ .)

The forces are not known explicitly. On the other hand, we know that at short distances there is a huge repulsive hard core part, and that the forces are attractive on medium distances, and they taper off very quickly after that. Moreover, protons interact via a Coulomb force.

That the forces are unknown is of course a huge problem. There are several approaches, such as fitting experimental scattering data to phenomenological models. More recently, direct calculations from QCD also has given some model potentials.

Another big problem is that the nucleus is *self-bound*. There are no external forces that bind the system. Indeed, the Hamiltonian is translation invariant, and thus has a purely continuous spectrum. This is remedied by writing

$$\hat{T} = \hat{T}_{\text{CoM}} + \hat{T}_{\text{int}}, \quad (2.173)$$

with

$$\hat{T}_{\text{CoM}} = \frac{1}{A} \vec{P}^2, \quad \hat{P} = \sum_i \vec{p}_i \quad (2.174)$$

being the centre-of-mass kinetic energy, and with

$$\hat{T}_{\text{int}} = \frac{1}{A} \sum_{i<j}^A (\vec{p}_i - \vec{p}_j)^2 \quad (2.175)$$

being the internal kinetic energy. We note that the *internal Hamiltonian*

$$\hat{H}_{\text{int}} = \hat{H} - \hat{T}_{\text{int}} \quad (2.176)$$

only depends on relative coordinates and momenta. The internal Hamiltonian has a discrete spectrum. The total energy is  $E = E_{\text{CoM}} + E_{\text{int}}$ , depending continuously on the CoM kinetic energy  $E_{\text{CoM}}$ .

[More to come later. References. Graphs.]

## 2.4.4 Exercises

**Exercise 2.28.** Consider a system consisting of  $N = 2$  electrons. Show that the wavefunction can be factorized as

$$\Psi(x_1, x_2) = \psi(\vec{r}_1, \vec{r}_2) \xi(s_1, s_2)$$

where either  $\psi$  is antisymmetric and  $\xi$  is symmetric or vice versa. Compute the eigenvalue of the total spin in each case.  $\triangle$

**Exercise 2.29.** Generalize the one-dimensional quantum dot model to three dimensions. You are allowed to set  $\delta = 0$  in this case.  $\triangle$

**Exercise 2.30.** Prove Eq. (2.173) for the nuclear manybody problem.  $\triangle$

## Chapter 3

# Hartree–Fock theory part I

### 3.1 The Hartree–Fock equations

Suggested reading for this section: Szabo/Ostlund Ch. 13, Harris/Monkhorst/Freeman Ch. 3, Gross/Runge/Heinonen Ch. 7.

One of the earliest and most successful approximation methods for many-fermion systems was the *Hartree–Fock method* (HF method). In Hartree–Fock theory we make an Ansatz, that the wavefunction is a single Slater determinant. The single-particle functions are the primary unknowns to be optimized using the variational principle.

Let  $\hat{H} = \hat{H}_0 + \hat{W}$  be the system Hamiltonian, with

$$\hat{H}_0 = \sum_i^N \hat{h}(i).$$

The ground-state of  $\hat{H}_0$  alone is a Slater determinant. The idea is that if  $\hat{W}$  is somehow “small”, it is reasonable that *some* Slater determinant will be a good approximation to the true ground-state wavefunction. As we will see, each particle moves independently of the others, experiencing a sort of average potential  $\hat{v}^{\text{HF}}$  set up by the other particles. Hartree–Fock theory is therefore called a *mean-field theory*, or an *independent-particle model*.

Thus, our wavefunction is a Slater determinant defined by a collection of  $N$  single-particle functions  $\phi_i$ , which we may take to be orthonormal,

$$|\Phi\rangle = |\phi_1\phi_2\cdots\phi_N\rangle, \quad \langle\phi_i|\phi_j\rangle = \delta_{ij}. \quad (3.1)$$

Note carefully, that a single-particle basis is not given – it is to be found.

The expectation value of the Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{W}$  now reads (recalling that  $\langle\Phi|\Phi\rangle = 1$ )

$$\langle\Phi|\hat{H}|\Phi\rangle = \sum_i \langle\phi_i|\hat{h}|\phi_i\rangle + \frac{1}{2} \sum_{ij} \langle\phi_i\phi_j|\hat{w}|\phi_i\phi_j - \phi_j\phi_i\rangle \quad (3.2)$$

as obtained from exercise 3.1. Here, the matrix elements of  $\hat{w}$  are *not* antisymmetrized, even if the notation suggests that  $\langle\phi_i\phi_j|$ , etc, are Slater determinants. Thus,

$$\langle\phi_i\phi_j|\hat{w}|\phi_k\phi_l\rangle \equiv \iint \overline{\phi_i(1)\phi_j(2)} \hat{w}(1,2) \phi_k(1)\phi_l(2) d1d2. \quad (3.3)$$

We recall the symmetry properties  $\langle \phi_p \phi_q | \hat{w} | \phi_r \phi_s \rangle = \langle \phi_q \phi_p | \hat{w} | \phi_s \phi_r \rangle$ . The task is now to minimize this energy  $\langle \Phi | \hat{H} | \Phi \rangle$  subject to the constraint that the  $\phi_i$  are orthonormalized,

$$\langle \phi_i | \phi_j \rangle = \delta_{ij}. \quad (3.4)$$

When a minimum is found, we denote the solution by  $|\Phi_{\text{HF}}\rangle$ , the Hartree–Fock state.

The constraints constitute a complication that we want to get rid of. We therefore introduce *Lagrange multipliers*, one for each constraint, giving a Lagrangian functional

$$\begin{aligned} \mathcal{L}[\phi_1, \dots, \phi_N, \lambda] &= \langle \Phi | \hat{H} | \Phi \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \\ &= \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_j \rangle - \delta_{ij}) \end{aligned} \quad (3.5)$$

Computing an extremum for the constrained problem is equivalent to an *unconstrained* extremalization of  $\mathcal{L}$  with respect to the  $\phi_i$  and the Lagrange multipliers.

Due to symmetry of the constraints, the Lagrange multiplier matrix  $\lambda$  can be assumed to be Hermitian<sup>1</sup>.

A word on a special notation. We define a single-particle function

$$\langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle \in L_1^2 \quad (3.6)$$

as the function obtained by integrating only over the second particle in the matrix element, viz,

$$\langle x_1 | \langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle (x_1) \equiv \int \phi_1(x_2)^* [w(x_1, x_2) \phi_2(x_1) \phi_3(x_2)] dx_2. \quad (3.7)$$

The inner product with any single-particle function  $\chi$  is

$$\langle \chi | \langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \iint \chi(x_1)^* \phi_1(x_2)^* [w(x_1, x_2) \phi_2(x_1) \phi_3(x_2)] dx_1 dx_2 = \langle \chi \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle, \quad (3.8)$$

i.e., the full two-particle integral. Thus, the dot represents an “unused slot” in the two-particle matrix element.

We can expand the function in any orthonormal single-particle basis  $\{\chi_p\} \subset L_1^2$ ,

$$\langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \sum_p |\chi_p\rangle \langle \chi_p | \langle \cdot | \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle = \sum_p |\chi_p\rangle \langle \chi_p \phi_1 | \hat{w} | \phi_2 \phi_3 \rangle, \quad (3.9)$$

i.e., a linear combination of two-particle matrix elements. This notation will be useful when we now state and prove our result:

**Theorem 3.1** (Hartree–Fock equations). *Let  $\{\phi_i\}$  be  $N$  orthonormal single-particle functions. Suppose the Slater determinant  $|\Phi_{\text{HF}}\rangle = |\phi_1 \cdots \phi_N\rangle$  is an extremum of the Hartree–Fock energy.*

*Then, the following non-linear system of equations is satisfied:*

$$\hat{f}(\phi_1, \dots, \phi_N) |\phi_i\rangle = \sum_j |\phi_j\rangle \lambda_{ji}, \quad (3.10)$$

where  $\lambda = [\lambda_{ij}]$  is a (Hermitian) matrix.

<sup>1</sup>To see this, assume that  $a_{ji}$  is a matrix which is not assumed to be Hermitian. Note that the expression  $g_{ij} = \langle \phi_i | \phi_j \rangle - \delta_{ij}$  satisfies  $g_{ij}^* = g_{ji}$ . Thus,  $\sum_{ij} a_{ji} g_{ij} = \sum_{ij} a_{ji} g_{ji}^* = \sum_{ij} a_{ij} g_{ij}^* = (\sum_{ij} a_{ij}^* g_{ij})^*$ . This gives  $\sum_{ij} a_{ji} g_{ij} = \frac{1}{2} \sum_{ij} (a_{ji} + a_{ij}^*) g_{ij}$ . Take  $\lambda_{ji} = a_{ji} + a_{ij}^*$ .

Here, the Fock operator  $\hat{f}$  is given by

$$\hat{f}(\phi_1, \dots, \phi_N) \equiv \hat{h} + \hat{v}^{HF}(\phi_1, \dots, \phi_N), \quad (3.11)$$

with

$$\hat{v}^{HF} |\psi\rangle \equiv (\hat{v}^{direct} - \hat{v}^{exchange}) |\psi\rangle = \sum_j \langle \cdot \phi_j | \hat{w} | \psi \phi_j \rangle - \sum_j \langle \cdot \phi_j | \hat{w} | \phi_j \psi \rangle. \quad (3.12)$$

The converse is also true, any solution to Eqs. (3.10) is an extremal point to the Hartree–Fock energy.

Moreover, by a unitary rotation of the  $\{\phi_i\}$ ,  $\lambda$  can be assumed to be diagonal, i.e., there exists a unitary matrix  $U = [U_{ij}]$  such that  $\tilde{\phi}_i = \sum_j \phi_j U_{ji}$  solves the equation

$$\hat{f}(\tilde{\phi}_1, \dots, \tilde{\phi}_N) |\tilde{\phi}_i\rangle = \epsilon_i |\tilde{\phi}_i\rangle. \quad (3.13)$$

The equations (3.13) are referred to as “the canonical Hartree–Fock equations”, while Eq. (3.10) are the “non-canonical Hartree–Fock equations”. The operator  $\hat{f}$  in Eq. (3.11) is “the Fock operator”, and  $\hat{v}^{direct}$  and  $\hat{v}^{exchange}$  are the direct- and exchange potentials, respectively.

*Proof.* In the language of Sec. A.1, we need to show that the directional derivative of the Lagrangian vanishes.

We first note that  $\lambda_{ji}$  can be treated separately:  $\partial \mathcal{L} / \partial \lambda_{ji} = \langle \phi_i | \phi_j \rangle - \delta_{ij}$ , the constraint. These equations are ensured fulfilled in the end by finding solutions  $\phi_i$  that are in fact orthonormal. We thus only compute the directional derivatives with respect to variations of the  $\phi_i$ . In fact, we restrict our attention to the derivatives with respect to the *complex conjugates* of  $\phi_i$ , since (a) we can treat  $\phi_i$  and  $\phi_i^*$  as independent variables, and since (b) variation with respect to the non-conjugated SPFs yield the same equations, but complex conjugated.

Choose a  $k \in \{1, \dots, N\}$ . We are going to compute the directional derivative with respect to changes in the function  $\phi_k^*$  only, leaving the other fixed. This turns out to be sufficient to find all the equations. Thus, let  $\epsilon$  be small, and let  $\eta$  be a normalized single-particle function. We write

$$\delta \phi_k = \epsilon \eta.$$

The other functions are fixed,  $\delta \phi_l = 0$  for  $l \neq k$ . Define the function

$$f(\epsilon) = \mathcal{L}(\phi_1, \dots, \phi_k + \epsilon \eta, \dots, \phi_N, \lambda), \quad (3.14)$$

To first order in  $\epsilon$ ,

$$f(\epsilon) = f(0) + \epsilon f'(0) + O(\epsilon^2), \quad (3.15)$$

and we for an extremal point of  $\mathcal{L}$ , we must have that for any  $\eta$ ,  $f'(0) = 0$ . In the language of Sec. A.1, the directional derivative of  $\mathcal{L}$  at  $\{\phi_i\}_{i=1}^N$  in the direction  $\eta$  (for  $\phi_k$ , the others are fixed) vanishes.

We compute the Taylor expansion of  $f(\epsilon)$  by direct computation of the perturbed Lagrangian, recalling that we only vary  $\phi_k^*$ :

$$\begin{aligned} f(\epsilon) &= \sum_i \langle \phi_i + \delta_{ki} \epsilon \eta | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle (\phi_i + \delta_{ki} \epsilon \eta)(\phi_j + \delta_{kj} \epsilon \eta) | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle \\ &\quad - \sum_{ij} \lambda_{ji} (\langle \phi_i + \delta_{ki} \epsilon \eta | \phi_j \rangle - \delta_{ij}) \end{aligned} \quad (3.16)$$



We now write out the matrix elements, but keep only terms up to first order in  $\epsilon$ . This gives

$$\begin{aligned}
f(\epsilon) = & \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle + \epsilon \langle \eta | \hat{h} | \phi_k \rangle \\
& + \frac{1}{2} \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle + \frac{1}{2} \epsilon \sum_i \langle \phi_i \eta | \hat{w} | \phi_i \phi_k \rangle - \frac{1}{2} \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle - \frac{1}{2} \epsilon \sum_i \langle \phi_i \eta | \hat{w} | \phi_k \phi_i \rangle \\
& - \sum_j \lambda_{jk} \epsilon \langle \eta | \phi_j \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_i \rangle - \delta_{ij}) + O(\epsilon^2)
\end{aligned} \tag{3.17}$$

We now use the symmetry property of the matrix elements of  $\hat{w}$ . This gives, for example,

$$\frac{1}{2} \sum_i \langle \phi_i \eta | \hat{w} | \phi_i \phi_k \rangle = \frac{1}{2} \sum_i \langle \eta \phi_i | \hat{w} | \phi_k \phi_i \rangle = \frac{1}{2} \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle. \tag{3.18}$$

This gives a simplification of  $f(\epsilon)$ , and we regroup:

$$\begin{aligned}
f(\epsilon) = & \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle - \sum_{ij} \lambda_{ji} (\langle \phi_i | \phi_i \rangle - \delta_{ij}) \\
& + \epsilon \langle \eta | \hat{h} | \phi_k \rangle + \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle - \epsilon \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle \\
& - \epsilon \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle + O(\epsilon^2)
\end{aligned} \tag{3.19}$$

We recognize that the zeroth order term is just  $f(0) = \mathcal{L}(\phi_1, \dots, \phi_N, \lambda)$ . We read off  $f'(\epsilon)$ , and obtain the directional derivative, and hence the equation

$$0 = \langle \eta | \hat{h} | \phi_k \rangle + \sum_j \langle \eta \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \eta \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} \langle \eta | \phi_j \rangle, \tag{3.20}$$

which must be valid for all choices of the function  $\eta$ . The only possibility for this to be true, is when

$$0 = \hat{h} | \phi_k \rangle + \sum_j \langle \cdot | \phi_j | \hat{w} | \phi_k \phi_j \rangle - \sum_j \langle \cdot | \phi_j | \hat{w} | \phi_j \phi_k \rangle - \sum_j \lambda_{jk} | \phi_j \rangle. \tag{3.21}$$

Thus, we have proven that the  $\phi_i$  solve the non-canonical HF equations if and only if  $|\Phi\rangle$  is an extremum of the HF energy.

We now show that we can diagonalize  $\lambda$ , producing the canonical HF equation. Thus, we replace  $\lambda$  it with a diagonal matrix with diagonal elements  $\epsilon_k$  (not to be confused with the small parameter  $\epsilon$  above, which we now are done with.)

From Exercise 2.7, we know that the determinant  $|\Phi\rangle$  is invariant (up to an irrelevant phase) under a unitary mixing of the single-particle functions, i.e, if we let

$$\tilde{\phi}_k = \sum_j \phi_j U_{jk} \tag{3.22}$$

with  $U$  a unitary matrix, then  $|\tilde{\Phi}\rangle = \det(U) |\Phi\rangle$ , i.e., the same state, and clearly the energy must be the same too.

As argued,  $\lambda_{ij} = \lambda_{ji}^*$  can be assumed Hermitian. Select therefore  $U$  such that  $\lambda = UEU^H$ , with  $E_{jk} = \delta_{jk}\epsilon_k$  the elements of a diagonal matrix (the eigenvalues of  $\lambda$ ):

$$\lambda_{ji} = \sum_{\ell} U_{j\ell} U_{i\ell}^*. \quad (3.23)$$

Let  $|r_i\rangle$  be the right-hand side of Eq. (3.21), and consider

$$\sum_k |r_k\rangle U_{ki} = 0. \quad (3.24)$$

Since  $U$  is unitary, Eq. (3.21) is satisfied for all  $k$  if and only if Eq. (3.24) is satisfied for all  $i$ . Computing the sum in Eq. (3.24) (see Exercise 3.2) we obtain

$$\hat{h} |\tilde{\phi}_i\rangle + \sum_j [\langle \cdot \tilde{\phi}_j | \hat{w} | \tilde{\phi}_i \tilde{\phi}_j \rangle - \langle \cdot \tilde{\phi}_j | \hat{w} | \tilde{\phi}_j \tilde{\phi}_i \rangle] - \epsilon_i |\tilde{\phi}_i\rangle = 0. \quad (3.25)$$

This must hold for all  $i = 1, \dots, N$  simultaneously.

With the definitions of  $\hat{v}^{\text{direct}}$  and  $\hat{v}^{\text{exchange}}$  in the theorem formulation, we are finished.  $\square$

The theorem does not guarantee that the solutions to the HF equations correspond to a the actual HF solution, i.e., a global minimum, or even a local minimum. It could well be a saddle point.

### 3.1.1 Exercises

**Exercise 3.1.** Prove Eq. (7.58), i.e., for a Slater determinant with orthonormal SPFs  $\phi_i$ ,  $i = 1, 2, \dots, N$ , we have

$$\langle \Phi | \hat{H} | \Phi \rangle = \sum_i \langle \phi_i | \hat{h} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle.$$

You may solve this problem by elementary application of the fundamental anticommutator.

Explain why this expectation value is invariant under a unitary mixing of the SPFs,  $\tilde{\phi}_i = \sum_j \phi_j U_{ji}$ , where  $U$  is a unitary matrix.  $\triangle$

**Exercise 3.2.** In this exercise, we fill in the details between Eq. (3.24) and Eq. (3.25) in the proof of Theorem 3.1.

a) Verify that

$$\sum_k U_{ki} \hat{h} |\phi_k\rangle = \hat{h} |\tilde{\phi}_i\rangle. \quad (3.26)$$

b) Next, show that

$$\sum_k U_{ki} \sum_j \lambda_{jk} |\phi_j\rangle = \epsilon_i |\tilde{\phi}_i\rangle. \quad (3.27)$$

c) As an intermediate calculation, verify that

$$|\phi_i\rangle = \sum_k U_{ki}^H |\tilde{\phi}_k\rangle = \sum_k U_{ik}^* |\tilde{\phi}_k\rangle. \quad (3.28)$$

d) Show that

$$\sum_k U_{ki} \sum_j \langle \cdot \phi_j | \hat{w} | \phi_k \phi_j \rangle = \sum_j \langle \cdot \tilde{\phi}_j | \hat{w} | \tilde{\phi}_i \tilde{\phi}_j \rangle. \quad (3.29)$$

You may do the transformations of the various  $\phi_\ell$  into  $\tilde{\phi}_\ell$  using c), or use Exercise 3.5.

e) Show that

$$\sum_k U_{ki} \sum_j \langle \cdot \phi_j | \hat{w} | \phi_j \phi_k \rangle = \sum_j \langle \cdot \tilde{\phi}_j | \hat{w} | \tilde{\phi}_j \tilde{\phi}_i \rangle. \quad (3.30)$$

f) Gather the results of a), b), d), and e), to show that Eq. (3.24) becomes Eq. (3.25).

△

**Exercise 3.3.** We now show an interesting relation for the Hartree–Fock energy. It is tempting to assume that  $E_{\text{HF}} = \sum_i \epsilon_i$ . However, this is not the case.

Assume that a solution  $(\phi_i, \epsilon_i)$ ,  $i = 1, \dots, N$ , to the canonical Hartree–Fock equations have been found. Show that the Hartree–Fock energy can be written

$$E_{\text{HF}} = \sum_i \epsilon_i - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_i | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle. \quad (3.31)$$

Explain why, in the non-canonical case, we have

$$E_{\text{HF}} = \sum_i \lambda_{ii} - \frac{1}{2} \sum_{ij} \langle \phi_i \phi_i | \hat{w} | \phi_i \phi_j - \phi_j \phi_i \rangle. \quad (3.32)$$

△

## 3.2 The direct and exchange potentials

Let us consider the unfamiliar operators  $\hat{v}^{\text{direct}}$  and  $\hat{v}^{\text{exchange}}$  in some detail. To this end, suppose that the two-body operator is a local potential  $\hat{w}(x_1, x_2)$ , such as the Coulomb potential in electronic systems,

$$\hat{w}_{\text{Coul}}(x_1, x_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|}, \quad x_i = (\vec{r}_i, \sigma). \quad (3.33)$$

The operator  $\hat{v}^{\text{direct}}$  is a one-body operator. When acting on a one-body function  $|\psi\rangle$  it produces a new one-body function, which at  $x_1$  takes the value

$$\begin{aligned} \langle x_1 | (\hat{v}^{\text{direct}} |\psi\rangle) &= \sum_j \langle x_1 | \langle \cdot | \phi_j | \hat{w} | \psi \phi_j \rangle = \sum_j \int \phi_j^*(x_2) w(x_1, x_2) \phi_j(x_2) \psi(x_1) dx_2 \\ &= \left[ \int \sum_j |\phi_j(x_2)|^2 w(x_1, x_2) dx_2 \right] \psi(x_1) \equiv v^{\text{direct}}(x_1) \psi(x_1). \end{aligned} \quad (3.34)$$

Thus,  $\hat{v}^{\text{direct}}$  is a *local potential*, given by a sort of average of  $w(x_2, x_1)$  over  $x_2$ , weighted by  $\rho(x) \equiv \sum_j |\phi_j(x)|^2$ , giving a “mean-field potential”.

The operator  $\hat{v}^{\text{exchange}}$  is, however, *non-local*: the value  $\langle x_1 | (\hat{v}^{\text{exchange}} |\psi\rangle)$  depends on  $\psi(x_2)$  in every point  $x_2$ . To see this, we compute

$$\langle x_1 | (\hat{v}^{\text{exchange}} |\psi\rangle) = \sum_j \langle x_1 | \langle \cdot | \phi_j | \hat{w} | \phi_j \psi \rangle = \sum_j \int \phi_j^*(x_2) w(x_1, x_2) \psi(x_2) \phi_j(x_1) dx_2. \quad (3.35)$$

The operator  $\hat{v}^{\text{exchange}}$  is still *linear* when acting on  $|\psi\rangle$ , it is just not interpretable as a local potential.

If we introduce the *reduced one-particle density matrix*  $\gamma(x, x')$  as

$$\gamma(x, x') = \sum_j \phi_j(x) \phi_j(x')^*, \quad (3.36)$$

we can express

$$\langle x_1 | \hat{v}^{\text{direct}} |\psi\rangle = \int \gamma(x_2, x_2) w(x_1, x_2) dx_2 \psi(x_1). \quad (3.37)$$

$$\langle x_1 | \hat{v}^{\text{exchange}} |\psi\rangle = \int \gamma(x_1, x_2) w(x_1, x_2) \psi(x_2) dx_2. \quad (3.38)$$

The reduced density matrix  $\gamma$  is a useful concept in Hartree–Fock theory.

In the proof of the HF equations, we first found an equation whose solutions were not eigenfunctions, Eq. (3.21). However, by forming a particular linear combination, the equation was brought on eigenvalue form, Eq. (3.13). We realized that the HF single-particle functions were not unique; any unitary transformation among the orbitals produces the *same*  $|\Phi_{\text{HF}}\rangle$ .

### 3.3 Augmenting the HF single-particle functions to a basis

We only used the  $N$  first eigenvectors of the Fock operator  $\hat{f} = \hat{h} + \hat{v}^{\text{HF}}$  to construct our HF wavefunction. But when these have been found,  $\hat{f} = \hat{f}(\phi_1, \dots, \phi_N)$  is a fixed Hermitian operator (see Exercise 3.4), and we can in principle<sup>2</sup> find a complete basis of eigenvectors of  $\hat{f}$ ,

$$\{\phi_p\} = \{\phi_i\} \cup \{\phi_a\}. \quad (3.39)$$

We here introduce the common convention that different indexed names implies different SPF index “types”: indices  $i, j, k, \dots$  imply occupied indices in the HF Slater determinant, while indices  $a, b, c, \dots$  imply unoccupied indices. Indices  $p, q, r, \dots$  imply general indices.

The HF single-particle basis is very useful for more involved manybody treatments, such as CI calculations, perturbation theory, and coupled-cluster (CC) theory. The HF single-particle functions can be used as any other single-particle basis, to set up Slater determinants, creation and annihilation operators, normal-ordered operators, etc. Each HF function  $\phi_p$  is associated with a  $c_p^\dagger$ .

This rather central, that we write it up as a definition that we can refer to later:

**Definition 3.1** (Canonical HF equations, HF basis). *For a given two-body Hamiltonian*

$$\hat{H} = \sum_{i=1}^N \hat{h}(i) + \sum_{i<j}^N \hat{w}(i, j), \quad (3.40)$$

*The equation*

$$\hat{f}(\phi_1, \dots, \phi_N) |\phi_P\rangle = \epsilon_P |\phi_P\rangle. \quad (3.41)$$

*with the Fock operator*

$$\hat{f}(\phi_1, \dots, \phi_N) = \hat{h} + \hat{v}^{\text{direct}} - \hat{v}^{\text{exchange}}, \quad (3.42)$$

*is referred to as the canonical Hartree–Fock equations, and the solutions are called the canonical single-particle functions  $\phi_p$ .*

*The first  $N$  HF single-particle functions  $\phi_i$  are often called occupied, while the rest,  $\phi_a$ , are often called virtual/unoccupied single-particle functions.*

Starting from the HF determinant, often called the “reference determinant”, we can make all other Slater determinant basis functions by successively replacing one occupied index  $i$  with a virtual index  $a$ . For example,

$$|\Phi_i^a\rangle \equiv c_a^\dagger c_i |\Phi\rangle \quad (3.43)$$

the index  $i$  is replaced by  $a$ . This is called a *singly excited determinant*. Similarly, two replacements lead to a *doubly excited determinant*,

$$|\Phi_{ij}^{ab}\rangle \equiv c_a^\dagger c_i c_b^\dagger c_j |\Phi\rangle. \quad (3.44)$$

Note that  $|\Phi_{ij}^{ab}\rangle = |\Phi_{ji}^{ba}\rangle = -|\Phi_{ij}^{ba}\rangle$  etc. We can continue, reaching up to  $N$  replacements,

$$|\Phi_{i_1 i_2 \dots i_N}^{a_1 a_2 \dots a_N}\rangle \equiv c_{a_1}^\dagger c_{i_1} \dots c_{a_N}^\dagger c_{i_N} |\Phi\rangle. \quad (3.45)$$

Why is the HF basis particularly useful? The idea is that while  $|\Phi\rangle = |\Phi_{\text{HF}}\rangle$  may be a good starting point, it is not a sufficiently good approximation. Thus, we wish to introduce a correction, i.e.,

$$|\Psi\rangle \approx |\Phi\rangle + |\delta\Psi\rangle, \quad (3.46)$$

---

<sup>2</sup>It happens that  $\hat{f}$  has a continuous spectrum, so our statement must really be limited to finite-dimensional one-particle spaces for strict validity.

where

$$|\delta\Psi\rangle = \sum_{IA} C_i^a |\Phi_i^a\rangle + \frac{1}{4} \sum_{ijab} C_{ij}^{ab} |\Phi_{ij}^{ab}\rangle + \dots \quad (3.47)$$

(Here, we have introduced a factor  $1/4$  to account for over-counting the Slater determinants.) The idea is that the *excitation amplitudes*  $C_i^a$ ,  $C_{ij}^{ab}$ , etc, should be small, and devise schemes based on this.

We also have Brillouin's Theorem, which leads to some simplifications in many cases:

**Theorem 3.2** (Brillouin's Theorem). *Using HF single-particle functions (canonical or non-canonical), the matrix element of  $\hat{H}$  between the reference and singly excited determinants vanish,*

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = 0. \quad (3.48)$$

*Proof.* See exercise 3.6. □

It is convenient to write the Hamiltonian on the following form

$$\hat{H} = \hat{H}_0 + \hat{W} = \hat{F} + \hat{U}, \quad (3.49)$$

where the second-quantized Fock operator is given by

$$\hat{F} = \sum_{i=1}^N \hat{f}(i) = \hat{H}_0 + \hat{V}^{\text{HF}}, \quad (3.50)$$

and where the *fluctuation potential* is given by

$$\hat{U} = \hat{W} - \hat{V}^{\text{HF}}. \quad (3.51)$$

Here,  $\hat{V}^{\text{HF}} = \hat{V}^{\text{direct}} - \hat{V}^{\text{exchange}}$ , with

$$\hat{V}^{\text{direct}} = \sum_i \hat{v}^{\text{direct}}(i), \quad \hat{V}^{\text{exchange}} = \sum_i \hat{v}^{\text{exchange}}(i). \quad (3.52)$$

An expression for the direct potential operator matrix element is easily obtained:

$$\langle \phi_Q | \hat{v}^{\text{direct}} | \phi_P \rangle = \sum_I \langle \phi_I \phi_Q | \hat{w} | \phi_I \phi_P \rangle, \quad (3.53)$$

with non-antisymmetric matrix elements. Thus,

$$\hat{V}^{\text{direct}} = \sum_{PQ} \left( \sum_I \langle \phi_I \phi_Q | \hat{w} | \phi_I \phi_P \rangle \right) c_Q^\dagger c_P. \quad (3.54)$$

Similarly, for the exchange potential we get

$$\hat{V}^{\text{exchange}} = \sum_{PQ} \left( \sum_I \langle \phi_I \phi_Q | \hat{w} | \phi_P \phi_I \rangle \right) c_Q^\dagger c_P. \quad (3.55)$$

In total, we have

$$\hat{V}^{\text{HF}} = \sum_{PQ} \left( \sum_I w_{PI}^{QI} \right) c_Q^\dagger c_P, \quad w_{RS}^{PQ} \equiv \langle QP | \hat{w} | RS \rangle_{AS}. \quad (3.56)$$

Note that we use antisymmetrized matrix elements. This results in the Fock operator and fluctuation potential

$$\hat{F} = \sum_{PQ} \left( h_P^Q + \sum_I w_{PI}^{QI} \right) c_Q^\dagger c_P \quad (3.57)$$

$$\hat{U} = \hat{W} - \sum_{PQ} \left( \sum_I w_{PI}^{QI} \right) c_q^\dagger c_p. \quad (3.58)$$

We write

$$\hat{F} = \sum_{QP} f_P^Q c_Q^\dagger c_P, \quad f_P^Q = \langle \phi_Q | \hat{F} | \phi_P \rangle = h_P^Q + \sum_I w_{PI}^{QI}. \quad (3.59)$$

So far we have *not used* that the  $\phi_P$  are HF single-particle functions. When we do assume this, we see from the non-canonical HF equations (3.10) that the matrix  $f = [f_P^Q]$  is block diagonal *if and only if* this equation is satisfied, i.e.,  $f_I^A = f_I^A = 0$ . Moreover, the Lagrange multipliers are given by  $\lambda_{JI} = f_J^I$ . In the case where the *canonical* HF equations are satisfied, we have  $f_Q^P = \delta_{PQ} \epsilon_Q$ .

### 3.3.1 Exercises

**Exercise 3.4.** Suppose the HF single-particle functions have been found, so that the Fock operator  $\hat{f}$  is a fixed operator. Prove that it is Hermitian, i.e., for any two single-particle functions  $\psi(x)$  and  $\psi'(x)$ ,

$$\langle \psi | \hat{f} | \psi' \rangle = [\langle \psi' | \hat{f} | \psi \rangle]^*.$$

△

**Exercise 3.5.** We show that the reduced one-particle density matrix is the same for canonical and non-canonical orbitals: Let  $U$  be a unitary matrix and define

$$\tilde{\phi}_i = \sum_j \phi_j U_{ji}. \quad (3.60)$$

Show that

$$\gamma(x, x') = \sum_j \tilde{\phi}_j(x) \tilde{\phi}_j(x')^*. \quad (3.61)$$

What can you conclude about  $\hat{v}^{\text{direct}}$  and  $\hat{v}^{\text{exchange}}$ , which are functions of  $\gamma$ ?

△

**Exercise 3.6.** In this exercise we prove Brillouin's Theorem: Using the HF basis, we have

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = 0.$$

a) Compute the matrix element

$$\langle \Phi_i^a | \hat{H}_0 | \Phi \rangle = \sum_{pq} h_q^p \langle \Phi_i^a | c_p^\dagger c_q | \Phi \rangle$$

Use elementary considerations of index replacements to determine all nonzero contributions.

**b)** Compute the matrix element

$$\langle \Phi_i^a | \hat{W} | \Phi \rangle = \frac{1}{2} \sum_{pqrs} w_{rs}^{pq} \langle \Phi_i^a | c_p^\dagger c_q^\dagger c_s c_r | \Phi \rangle$$

Use elementary considerations of index replacements to determine all nonzero contributions.

**c)** Add the results and use the (non-canonical) HF equation to show Brillouin's Theorem.

△



### 3.4 The Hartree–Fock equations in a given basis: the Roothan–Hall equations

How do we solve the HF equations (3.13)? In this section, we reformulate the HF equations relative to a fixed basis,  $\{\chi_p\}_{p=1}^L$ . For practical reasons, of course, the basis must have a finite size  $L$ . However, we do *not* assume that it is orthonormal. Thus, we have a *possibly non-diagonal overlap matrix*  $S$  of size  $L \times L$ ,

$$S_{pq} \equiv \langle \chi_p | \chi_q \rangle. \quad (3.62)$$

and we must have that  $S^{-1}$  exists since the  $\phi_p$  form a basis.

Such basis functions are common in quantum chemistry, where a non-orthogonal basis of *Gaussian functions centered on the atoms* is typically employed. See for example Szabo/Ostlund or Helgaker/Jørgensen/Olsen for details. For now, we just keep this remark as a motivation for not assuming orthogonality. In nuclear physics or solid state physics, orthogonal functions  $\chi_p$  are more typical.

We expand our HF functions as

$$|\phi_p\rangle = \sum_q |\chi_q\rangle U_{qp}, \quad (3.63)$$

where  $U$  is in general not a unitary matrix, since the basis is not orthogonal. (However, we have  $U^H S U = I$ , the identity matrix, see Exercise 3.7.) We notice that the *columns* of  $U$  are the basis expansions of each  $\phi_p$ . We write  $u_p$  for column number  $p$ ,  $|\phi_p\rangle = \sum_q |\chi_q\rangle (u_p)_q$ .

The reduced density matrix becomes

$$\gamma(x, x') = \sum_i \langle x | \phi_i \rangle \langle \phi_i | x' \rangle = \sum_{pq} \sum_i U_{qi} |\chi_q\rangle \langle \chi_p | U_{pi}^* = \sum_{pq} \left( \sum_i U_{qi} U_{pi}^* \right) \langle x | \chi_q \rangle \langle \chi_p | x' \rangle = \sum_{pq} (U_{1:N} U_{1:N}^H)_{qp} \langle x | \chi_q \rangle \langle \chi_p | x' \rangle \quad (3.64)$$

and it makes sense to define

$$D = U_{1:N} U_{1:N}^H = \sum_i u_i u_i^H, \quad (3.65)$$

which we interpret as the reduced density matrix relative to the given basis  $\{\chi_p\}$ , depending on the  $N$  first columns of  $U$  only.

We now demonstrate how the canonical HF equations (3.41) can be written

$$F(D)U = S U \epsilon, \quad (3.66)$$

where

$$F_{pq} = \langle \chi_p | \hat{f}(\phi_1, \dots, \phi_N) | \chi_q \rangle \quad (3.67)$$

are the matrix elements of the Fock operator in the fixed basis, and where  $\epsilon = \text{diag}(\epsilon_1, \dots, \epsilon_L)$  is a diagonal matrix. Equation (3.41) is a nonlinear generalized eigenvalue problem.

Let us look at the matrix elements of  $f$ ,

$$F_{qp} = \langle \chi_q | \hat{f} | \chi_p \rangle = \langle \chi_q | \hat{h} | \chi_p \rangle + \langle \chi_q | \hat{v}^{\text{direct}} | \chi_q \rangle - \langle \chi_q | \hat{v}^{\text{exchange}} | \chi_p \rangle. \quad (3.68)$$

The direct term is

$$\begin{aligned} \langle \chi_q | \hat{v}^{\text{direct}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \chi_p \phi_j \rangle = \sum_{p'q'j} U_{jq'} U_{jp'}^* \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle. \end{aligned} \quad (3.69)$$

Correspondingly,

$$\begin{aligned}\langle \chi_q | \hat{v}^{\text{exchange}} | \chi_p \rangle &= \sum_j \langle \chi_q \phi_j | \hat{w} | \phi_j \chi_p \rangle = \sum_{p'q'j} U_{jq'} U_{jp'}^* \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle \\ &= \sum_{p'q'} D_{q'p'} \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle.\end{aligned}\quad (3.70)$$

We obtain

$$F_{qp} = \langle \chi_q | \hat{h} | \chi_p \rangle + \sum_{p'q'} D_{p'q'} (\langle \chi_q \chi_{q'} | \hat{w} | \chi_p \chi_{p'} \rangle - \langle \chi_q \chi_{q'} | \hat{w} | \chi_{p'} \chi_p \rangle). \quad (3.71)$$

Note that we have expressed  $F_{qp}$  in terms of *non-antisymmetric* matrix elements of  $\hat{w}$ .

Thus, projecting the LHS of the canonical HF equations onto the basis gives

$$\langle \chi_q | \hat{f} | \phi_p \rangle = \sum_{q'} \langle \chi_q | \hat{f} | \chi_{q'} \rangle U_{q'p} = \sum_{q'} F_{qq'} U_{q'p}, \quad \forall q, p. \quad (3.72)$$

The right-hand side gives the projection

$$\langle \chi_q | \phi_p \rangle \epsilon_p = \sum_{q'} \langle \chi_q | \chi_{q'} \rangle U_{q'p} \epsilon_p = \sum_{q'} S_{qq'} U_{q'p} \epsilon_i, \quad \forall q, p. \quad (3.73)$$

Gathering, we find

$$F(D)U = SU\epsilon, \quad (3.74)$$

and we are finished. This equation is called the Roothan–Hall equation.

In terms of each column, i.e., each  $\phi_p$ ,

$$F(D)u_p = \epsilon_p S u_p. \quad (3.75)$$

### 3.4.1 Exercises

**Exercise 3.7.** Prove that  $U^H S U = I$  (the identity matrix) by using  $\langle \phi_p | \phi_q \rangle = \delta_{pq}$  and

$$|\phi_p\rangle \sum_q S_{qp} |\chi_q\rangle, \quad S_{qp} = \langle \chi_q | \chi_p \rangle. \quad (3.76)$$

△

### 3.5 Self-consistent field iteration

How do we find self-consistent solutions of Eq. (3.75)? The standard approach is by self-consistent field interactions (SCF iterations), Finding hopefully better and better approximations  $u_i^{(k)}$ ,  $k = 1, 2, 3, \dots$ , to the canonical HF functions, starting from a well-selected initial guess  $u_i^{(0)}$ .

Let  $D^{(k)} = \sum_i u_i^{(k)} (u_i^{(k)})^H$  be the  $k$ 'th iteration's density matrix. Then, the basic SCF iteration is to compute a complete set of orthonormal vectors

$$F(D^{(k)})u_p^{(k+1)} = \epsilon_p^{(k+1)} S u_p^{(k+1)} \quad (3.77)$$

by numerical diagonalization, sorting the eigenvalues  $\epsilon_p^{(k+1)}$  in ascending order. Then,  $p = 1, \dots, N$  gives the next approximation to the HF eigenpairs  $(\phi_i, \epsilon_i)$ , while the next  $L - N$  form the additional canonical functions.

If the SCF iteration converges, it often converges to a solution that corresponds to the true HF minimum wavefunction. Sometimes it does not converge to the true solution, but is still useful. Sometimes it does not converge at all, and one needs to “fix” the SCF iteration.

In fact, the basic SCF iteration has very problematic convergence properties. The most common scheme today is the so-called *direct inversion in the iterative subspace* iteration (DIIS), but this is out of scope for the present course. Read more about DIIS in Helgaker/Jørgensen/Olsen [6], and see also <https://en.wikipedia.org/wiki/DIIS>.

### 3.6 RHF: Restricted Hartree–Fock [needs work] (optional)

[There is an unfortunate overlap between the notation for spin functions  $\chi_\alpha$  and the basis functions  $\chi_p$  in the previous section. Hopefully no confusion arises.]

We now discuss the *restricted Hartree–Fock (RHF) method* for electronic systems. Supporting material: Szabo and Ostlund.

Motivation:

Consider  $N$  electrons, which we assume to a first approximation do not interact among themselves, i.e., we neglect the inter-electron repulsion operator given by

$$\hat{w}(\vec{r}_1, \vec{r}_2) = \frac{1}{|\vec{r}_1 - \vec{r}_2|}, \quad (3.78)$$

in suitable units. The electrons are thus described by a one-body Hamiltonian  $\hat{H}_0 = \sum_i \hat{h}(i)$ ,

$$\hat{h}(\vec{r}) = -\frac{1}{2}\nabla^2 + v(\vec{r}), \quad (3.79)$$

where  $v(\vec{r})$  is an external electrostatic potential, such as the one set up by an atomic nucleus. The operator  $\hat{h}$  does not couple to electron spin, so that the single-particle eigenfunctions of  $\hat{h}$  separate,

$$\phi_P(\vec{r}, \sigma) = \varphi_p(\vec{r})\chi_\alpha(\sigma), \quad P = (p, \sigma), \quad (3.80)$$

where  $\alpha = \pm 1/2$  is the value of the projection of the electron spin along the  $z$ -axis. Also,  $\sigma = \pm 1$ , and  $\langle \chi_\alpha | \chi_\beta \rangle = \delta_{\alpha\beta}$ . The eigenvalue problem of  $\hat{h}(\vec{r})$  becomes

$$\hat{h}\varphi_p(\vec{r})\chi_\alpha(\sigma) = e_p\varphi_p(\vec{r})\chi_\alpha(\sigma), \quad \sigma = \pm 1. \quad (3.81)$$

where the eigenvalue  $e_p$  is seen to be doubly degenerate due to spin. The  $N$ -electron ground-state of  $\hat{H}_0$  is now given by the Slater determinant with the  $N$  first eigensolutions  $\phi_{(p,\sigma)}$  occupied. Assuming  $N$  even, we get

$$|\Phi\rangle = |\phi_{1,+}\phi_{1,-}\cdots\phi_{\frac{N}{2},+}\phi_{\frac{N}{2},-}\rangle. \quad (3.82)$$

A common notation is

$$|\Phi_{\text{RHF}}\rangle = |\varphi_1\bar{\varphi}_1\varphi_2\bar{\varphi}_2\cdots\varphi_{N/2}\bar{\varphi}_{N/2}\rangle \quad (3.83)$$

with the understanding that  $\varphi_p$  represents  $\phi_{p,+1/2}$  and  $\bar{\varphi}_p$  represents  $\phi_{p,-1/2}$ .

The idea of RHF is to assume that the exact ground-state has a similar structure. Thus, we do not optimize all the  $N$  single-particle functions freely, we assume that they form a set of doubly occupied orbitals. In RHF we therefore compute the HF single-particle functions by minimizing the energy under the assumption that  $|\Phi\rangle$  is on the form (3.83).

The HF energy is simplified because of the special case of single-particle functions on factorized spin-orbital form. Consider for example the matrix element

$$\langle \phi_{p,\alpha} | \hat{h} | \phi_{q,\beta} \rangle = \langle \chi_\alpha | \chi_\beta \rangle \int \varphi_p(\vec{r})^* \hat{h}(\vec{r}) \varphi_q(\vec{r}) d\vec{r} \equiv \delta_{\alpha\beta} (\varphi_p | \hat{h} | \varphi_q), \quad (3.84)$$

where we have introduced a special notation for the spatial matrix element. Similarly,

$$\langle \phi_{p\alpha}\phi_{q\beta} | \hat{w} | \phi_{r\gamma}\phi_{s\delta} \rangle = \langle \chi_\alpha | \chi_\gamma \rangle \langle \chi_\beta | \chi_\delta \rangle \iint \varphi_p(\vec{r}_1)^* \varphi_q(\vec{r}_2)^* \hat{w}(\vec{r}_1, \vec{r}_2) \varphi_r(\vec{r}_1) \varphi_s(\vec{r}_2) d\vec{r}_1 d\vec{r}_2 \equiv \delta_{\alpha\gamma} \delta_{\beta\delta} (\varphi_p \varphi_q | \hat{w} | \varphi_r \varphi_s), \quad (3.85)$$

where we also introduce a special notation to be used in the sequel.

We use Eqs. (3.84–3.85) and compute the energy of  $|\Phi\rangle$ :

$$\begin{aligned}
\langle \Phi | \hat{H} | \Phi \rangle &= \sum_{\alpha} \sum_{i=1}^{N/2} \langle \phi_{i\alpha} | \hat{h} | \phi_{i\alpha} \rangle + \frac{1}{2} \sum_{\alpha} \sum_{i=1}^{N/2} \sum_{\beta} \sum_{j=1}^{N/2} \langle \phi_{i\alpha} \phi_{j\beta} | \hat{w} | \phi_{i\alpha} \phi_{j\beta} - \phi_{j\beta} \phi_{i\alpha} \rangle \\
&= 2 \sum_{i=1}^{N/2} (\varphi_i | \hat{h} | \varphi_i) + 2 \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_i \varphi_j) - \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_j \varphi_i) \\
&\equiv 2 \sum_i (i | \hat{h} | i) + 2 \sum_{ij} [ij | ij], \quad [pq | rs] \equiv (\varphi_p \varphi_q | \hat{w} | \varphi_r \varphi_s) - \frac{1}{2} (\varphi_p \varphi_q | \hat{w} | \varphi_s \varphi_r).
\end{aligned} \tag{3.86}$$

Observe the factor 2 in front of the two first terms.

The RHF state is obtained by minimizing the energy with respect to orthonormal orbitals  $\varphi_i$ ,  $i = 1, \dots, N/2$ . We obtain the restricted HF equation.

**Theorem 3.3** (Restricted Hartree–Fock equations). *The orbitals of the minimizing RHF state  $|\Phi_{RHF}\rangle$  can be taken to satisfy the canonical RHF equations:*

$$\hat{f}(\gamma) \varphi_i(\vec{r}) = \epsilon_i \varphi_i(\vec{r}), \quad i = 1, \dots, N/2, \tag{3.87}$$

where the (RHF) Fock operator is given by

$$\hat{f}(\gamma) = \hat{h} + \tag{3.88}$$

and where the reduced density matrix is

$$\gamma(\vec{r}, \vec{r}') = 2 \sum_i \varphi_i(\vec{r}) \varphi_i(\vec{r}')^*. \tag{3.89}$$

The RHF energy is

$$E_{RHF} = 2 \sum_{i=1}^{N/2} \epsilon_i - 2 \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_i \varphi_j) + \sum_{ij}^{N/2} (\varphi_i \varphi_j | \hat{w} | \varphi_j \varphi_i) \tag{3.90}$$

*Proof.* (Optional reading.) Optimization of RHF energy, and RHF equations: Introducing Lagrange multipliers for the orthonormality constraints, we obtain a Lagrangian

$$\mathcal{L}[\varphi_1, \dots, \varphi_{N/2}, \lambda] = 2 \sum_i (\varphi_i | \hat{h} | \varphi_i) + 2 \sum_{ij} (\varphi_i \varphi_j | \hat{w} | \varphi_i \varphi_j) - \sum_{ij} (\varphi_i \varphi_j | \hat{w} | \varphi_j \varphi_i) - 2 \sum_{ij} \lambda_{ji} [(\varphi_i | \varphi_j) - \delta_{ij}], \tag{3.91}$$

where we have introduced Lagrange multipliers for the orthonormality constraints. The factor 2 in front of the constraint term is for convenience.

A procedure similar to the derivation of the HF equations (see Exercise 3.8) gives:

$$\hat{h} \varphi_i + 2 \sum_j (\cdot \varphi_j | \hat{w} | \varphi_i \varphi_j) - \sum_j (\cdot \varphi_j | \hat{w} | \varphi_j \varphi_i) - \sum_j \lambda_{ji} \varphi_j = 0. \tag{3.92}$$

A unitary transformation similar to the one for the HF equations allow us to replace  $\lambda$  by a diagonal matrix, finally obtaining

$$[\hat{h} + \hat{v}^{\text{Coulomb}} - \hat{v}^{\text{exchange}}] \varphi_i(\vec{r}) = \epsilon_i \varphi_i(\vec{r}), \quad i = 1, \dots, N/2, \tag{3.93}$$

with

$$\hat{v}^{\text{Coulomb}}(\vec{r}) = \int \gamma(\vec{r}', \vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} d\vec{r}' \tag{3.94}$$

being a local potential, and where

$$[\hat{v}^{\text{exchange}}\psi](\vec{r}) = \frac{1}{2} \int \gamma(\vec{r}', \vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} \psi(\vec{r}') d\vec{r}' \quad (3.95)$$

is a non-local potential. The reduced density matrix is

$$\gamma(\vec{r}, \vec{r}') \equiv 2 \sum_{j=1}^{N/2} \varphi_j(\vec{r}) \varphi_j(\vec{r}')^* \quad (3.96)$$

The proof of Eq. (3.90) is obtained by taking the inner product of Eq. (3.87) with  $\varphi_i$  and summing over  $i$ , then multiplying with 2.  $\square$

**Exercise 3.8.** In this exercise, we prove Theorem 3.3 (To be filled in.)

$\triangle$

### 3.7 UHF: Unrestricted Hartree–Fock theory [needs work] (optional)

Supporting material: Szabo and Ostlund.

The RHF model is usually a good approximation, but fails in some circumstances. The *unrestricted* Hartree–Fock model is an intermediate between the general HF model and the restricted HF model. In RHF space orbital  $i$  for both spins were required to be identical. In UHF we allow them to be different,

$$\phi_{i,\alpha}(\vec{r}, \sigma) = \varphi_i^\alpha(\vec{r})\chi_\alpha(\sigma). \quad (3.97)$$

Thus, the orbital carries a spin-index as well as a space index, compare with the RHF model. The UHF state can be written

$$|\Phi_{\text{UHF}}\rangle = |\varphi_1^{1/2} \bar{\varphi}_1^{-1/2} \varphi_2^{1/2} \bar{\varphi}_2^{-1/2} \dots \varphi_{N/2}^{1/2} \bar{\varphi}_{N/2}^{-1/2}\rangle, \quad (3.98)$$

compare with Eq. (??) Notice that the spin-orbitals are still orthogonal for different spins. Notice also that the general HF model is more general than UHF: there, each spin-orbital was not required to separate into a product of space and spin functions.

The UHF energy expectation value is (see Exercise 3.9)

$$E_{\text{UHF}} = \sum_{\alpha} \sum_{i=1}^{N/2} (\varphi_i^\alpha | \hat{h} | \varphi_i^\alpha) + \frac{1}{2} \sum_{\alpha} \sum_{\beta} \sum_{ij}^{N/2} (\varphi_i^\alpha \varphi_j^\beta | \hat{w} | \varphi_i^\alpha \varphi_j^\beta) - \frac{1}{2} \sum_{\alpha} \sum_{ij}^{N/2} (\varphi_i^\alpha \varphi_j^\alpha | \hat{w} | \varphi_j^\alpha \varphi_i^\alpha). \quad (3.99)$$

The variational UHF equations become

$$\hat{h} \varphi_i^\alpha(\vec{r}) + \sum_{\beta} \sum_j (\cdot \varphi_j^\beta | \hat{w} | \varphi_i^\alpha \varphi_j^\beta) - \sum_j (\cdot \varphi_j^\alpha | \hat{w} | \varphi_j^\alpha \varphi_i^\alpha) = \epsilon_i^\alpha \varphi_i^\alpha(\vec{r}), \quad (3.100)$$

where we note that each spin-orbital is not doubly degenerate anymore. We introduce the UHF Coulomb potential,

$$v^{\text{Coulomb}}(\vec{r}) = \int \sum_{j\beta} |\varphi_j^\beta(\vec{r}')|^2 \frac{1}{|\vec{r} - \vec{r}'|} d\vec{r}', \quad (3.101)$$

and the UHF exchange potential operator

$$[\hat{v}^{\alpha, \text{exchange}} \psi](\vec{r}) = \int \sum_j \varphi_j^\alpha(\vec{r}) \varphi_j^\alpha(\vec{r}')^* \psi(\vec{r}) \frac{1}{|\vec{r} - \vec{r}'|} d\vec{r}', \quad (3.102)$$

to obtain

$$[\hat{h} + \hat{v}^{\text{Coulomb}} - \hat{v}^{\alpha, \text{exchange}}] \phi_i^\alpha(\vec{r}) = \epsilon_i^\alpha \phi_i^\alpha(\vec{r}). \quad (3.103)$$

**Exercise 3.9.** Prove Eq. (3.99), by showing

$$\langle \Phi_{\text{UHF}} | \hat{H} | \Phi_{\text{UHF}} \rangle = E_{\text{UHF}}. \quad (3.104)$$

△

# Chapter 4

## Wick's Theorem and normal ordering

### 4.1 Wick's Theorem

Supporting material for this section is Shavitt/Bartlett Ch. 3, Gross/Runge/Heinonen Ch. 19.

#### 4.1.1 A motivation

We have in place several basic ingredients useful for computation: The description of wavefunctions in terms of creation and annihilation operators, and the operators as polynomials in creation- and annihilation operators. We also often need to compute *matrix elements* of operators between two states. For example, diagonal matrix elements like

$$\langle \Phi | \hat{H}_0 | \Phi \rangle = \sum_{qr} \langle q | \hat{h} | r \rangle \langle - | c_{p_N} c_{p_{N-1}} \cdots c_{p_1} c_q^\dagger c_r c_{p_1}^\dagger c_{p_2}^\dagger \cdots c_{p_N}^\dagger | - \rangle \quad (4.1)$$

with  $|\Phi\rangle = |p_1 p_2 \cdots p_N\rangle$ , and similarly

$$\langle \Phi | \hat{W} | \Phi \rangle = \frac{1}{4} \sum_{\substack{q_1 q_2 \\ r_1 r_2}} \langle q_1 q_2 | \hat{w} | r_1 r_2 \rangle_{AS} \langle - | c_{p_N} c_{p_{N-1}} \cdots c_{p_1} c_{q_1}^\dagger c_{q_2}^\dagger c_{r_2} c_{r_1} c_{p_1}^\dagger c_{p_2}^\dagger \cdots c_{p_N}^\dagger | - \rangle. \quad (4.2)$$

Notice that we used

$$|p_1 \cdots p_N\rangle = c_{p_1}^\dagger \cdots c_{p_N}^\dagger | - \rangle \quad (4.3)$$

and, by taking the adjoint,

$$\langle p_1 \cdots p_N | = \langle - | c_{p_N} \cdots c_{p_1}. \quad (4.4)$$

Observe that the order of the annihilation operators is the reverse of the order of the creation operators.

The number  $\langle - | c^{(\dagger)} c^{(\dagger)} \cdots c^{(\dagger)} | - \rangle$  is referred to a *vacuum expectation value*, and the problem of computing matrix elements is basically reduced to computing these.

Vacuum expectation values are also important for other computational methods we will encounter, like perturbation theory and coupled-cluster theory.

Let us consider an example, and compute a typical quantity occurring in the  $\hat{H}_0$  matrix element:

$$A = \langle p_1 p_2 | c_r^\dagger c_s | q_1 q_2 \rangle = \langle - | c_{p_2} c_{p_1} c_r^\dagger c_s c_{q_1}^\dagger c_{q_2}^\dagger | - \rangle. \quad (4.5)$$

Now, how are we going to approach this problem? Recall the anticommutation relations,

$$c_p c_q^\dagger + c_q^\dagger c_p = \delta_{pq} \quad (4.6)$$

$$c_p c_q + c_q c_p = 0 \quad (4.7)$$



and

$$c_p^\dagger c_q^\dagger + c_q^\dagger c_p^\dagger = 0. \quad (4.8)$$

So, we can “flip” two creation or annihilation operators adjacent to each other and compensate with a  $-$  sign. We can “flip” an annihilation and creation operator by a  $-$  sign, *but* we have to “pay a price” in the form of a Kronecker delta, an additional term. However, this additional term has two less creation and annihilation operators.

In this way, we can systematically move the annihilation operators *to the right*, and the creation operators *to the left*, possibly inserting Kronecker deltas and generating new terms with fewer operators. But when the annihilation operators are to the right they give zero contribution since  $c_p |-\rangle = 0$  for any annihilation operator.

Let us see this in practice, and first remove one pair of creation and annihilation operators:

$$\begin{aligned} A &= \langle - | c_{p_2} c_{p_1} c_r^\dagger c_s c_{q_1}^\dagger c_{q_2}^\dagger | - \rangle = \langle - | c_{p_2} c_{p_1} c_r^\dagger (\delta_{sq_1} - c_{q_1}^\dagger c_s) c_{q_2}^\dagger | - \rangle \\ &= \delta_{sq_1} \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_2}^\dagger | - \rangle - \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_1}^\dagger c_s c_{q_2}^\dagger | - \rangle \\ &= \delta_{sq_1} \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_2}^\dagger | - \rangle - \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_1}^\dagger (\delta_{sq_2} - c_{q_2}^\dagger c_s) | - \rangle \\ &= \delta_{sq_1} \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_2}^\dagger | - \rangle - \delta_{sq_2} \langle - | c_{p_2} c_{p_1} c_r^\dagger c_{q_1}^\dagger | - \rangle. \end{aligned} \quad (4.9)$$

We continue:

$$\begin{aligned} A &= \delta_{sq_1} \langle - | c_{p_2} (\delta_{p_1 r} - c_r^\dagger c_{p_1}) c_{q_2}^\dagger | - \rangle - \delta_{sq_2} \langle - | c_{p_2} (\delta_{p_1 r} - c_r^\dagger c_{p_1}) c_{q_1}^\dagger | - \rangle \\ &= \delta_{sq_1} \delta_{p_1 r} \langle - | c_{p_2} c_{q_2}^\dagger | - \rangle - \delta_{sq_1} \langle - | c_{p_2} c_r^\dagger c_{p_1} c_{q_2}^\dagger | - \rangle - (q_1 \leftrightarrow q_2). \end{aligned} \quad (4.10)$$

In the last equality, we have indicated that the remaining terms are generated from the previous ones by exchanging  $q_1$  and  $q_2$ .

Continuing,

$$A = \delta_{sq_1} \delta_{p_1 r} \langle - | c_{p_2} c_{q_2}^\dagger | - \rangle - \delta_{sq_1} \delta_{p_1 q_2} \langle - | c_{p_2} c_r^\dagger | - \rangle + \delta_{sq_1} \langle - | c_{p_2} c_r^\dagger c_{q_2}^\dagger c_{p_1} | - \rangle - (q_1 \leftrightarrow q_2). \quad (4.11)$$

Only the two first terms are non-vanishing, and we note, for example, that  $\langle - | c_{p_2} c_{q_2}^\dagger | - \rangle = \langle p_2 | q_2 \rangle = \delta_{p_2 q_2}$ . (We could also use the anticommutator once more.) This gives:

$$A = \delta_{sq_1} \delta_{p_1 r} \delta_{p_2 q_2} - \delta_{sq_1} \delta_{p_1 q_2} \delta_{p_2 r} - (p_1 \leftrightarrow q_2). \quad (4.12)$$

Yes, our life was made easier by introducing second-quantization. However, the matrix elements are still quite hard to compute. This is where *Wick's theorem* comes in, by giving a quicker way of performing the moving-of-annihilation-operators-to-the-right.

Observe that the vacuum expectation value is *basis independent*. The value only depends on the anticommutator relations, and these only depended on the orthonormality of  $\{\phi_p\}$ , and the fact that  $c_p |-\rangle = 0$ .

### 4.1.2 Vacuum expectation values

Consider the computation of a *vacuum expectation value* of a *string of creation and annihilation operators*:

$$\langle - | A_1 A_2 \cdots A_n | - \rangle, \quad (4.13)$$

where each of the  $A_i$  are one of the  $c_q^\dagger$  or  $c_q$ . For example, the overlap between two determinants is on this form:

$$\langle p_1 \cdots p_N | q_1 \cdots q_N \rangle = \langle - | c_{p_N} c_{p_{N-1}} \cdots c_{p_1} c_{q_1}^\dagger c_{q_2}^\dagger \cdots c_{q_N}^\dagger | - \rangle. \quad (4.14)$$

Another example is the matrix elements of an operator on second quantized form, say  $\hat{H}_0$ :

$$\langle p_1 \cdots p_N | \hat{H}_0 | q_1 \cdots q_N \rangle = \sum_{pq} \langle p | \hat{h} | q \rangle \langle p_1 \cdots p_N | c_p^\dagger c_q | q_1 \cdots q_N \rangle \sum_{pq} \langle p | \hat{h} | q \rangle \langle - | c_{p_N} \cdots c_{p_1} c_p^\dagger c_q c_{p_1}^\dagger \cdots c_{p_N}^\dagger | - \rangle. \quad (4.15)$$

The right-hand side is a linear combination of vacuum expectation values. So we see that having a straightforward way to compute Eq. (4.13) would be of great help.

Wick's Theorem is what we shall need.

### 4.1.3 Normal ordering and contractions

In this section, we denote a general string of  $n$  creation and annihilation operators by

$$A_1 A_2 \cdots A_n, \quad A_i \in \{c_p\} \cup \{c_p^\dagger\}. \quad (4.16)$$

We will state and prove *Wick's Theorem*, named after Gian Carlo Wick (1901–1992), Italian theoretical physicist [Phys. Rev. **80**, 268 (1950)]. Wick's theorem is based on two fundamental concepts, namely *normal ordering* and *contraction*.

**Definition 4.1** (Normal ordering). *The normal-ordered product form of an operator string  $A_1 A_2 \cdots A_n$  is defined as a rearrangement,*

$$\{A_1 A_2 \cdots A_n\} \equiv (-1)^{|\sigma|} A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(n)}, \quad (4.17)$$

where  $\sigma$  is a permutation such that all the creation operators in the operator string is to the left of all the annihilation operators, i.e.,

$$\{A_1 A_2 \cdots A_{n-1} A_n\} \equiv (-1)^{|\sigma|} [\text{creation operators}] \cdot [\text{annihilation operators}]. \quad (4.18)$$

*Normal-ordering of a linear combination of operator strings is defined by*

$$\{\alpha A_1 \cdots A_n + \beta B_1 \cdots B_m\} \equiv \alpha \{A_1 \cdots A_n\} + \beta \{B_1 \cdots B_m\}. \quad (4.19)$$

The permutation  $\sigma$  is in general not be unique, since we may permute the creation and annihilation operators separately without affecting the total expression. For example,

$$\{c_p c_q^\dagger c_r^\dagger c_s\} = c_q^\dagger c_r^\dagger c_p c_s = -c_r^\dagger c_q^\dagger c_p c_s = c_r^\dagger c_q^\dagger c_s c_p = -c_q^\dagger c_r^\dagger c_s c_p. \quad (4.20)$$

There are  $2 \times 2$  possible arrangements of the creation and annihilation operators that conform to the definition of the normal-ordered product in this case. See exercise ??

To find the permutation  $\sigma$  that normal-orders an operator product, it is usually simplest to count the number  $f$  of anticommutations necessary to achieve the rearrangement, and set  $(-1)^{|\sigma|} = (-1)^f$ .

Note that the string  $A_1 \cdots A_n \neq \{A_1 \cdots A_n\}$  in general, since by reordering creation and annihilation operators we neglect the extra terms arising from the Kronecker delta in the anti-commutator relation  $\{c_p, c_q^\dagger\} = \delta_{pq}$ .

Mathematical aside for the interested reader:  $\{\cdots\}$  is defined as a linear operator on the space of linear combinations of operator strings. The second-quantized formulas for  $\hat{H}_0$ ,  $\hat{W}$ , etc., are examples of such objects. This space of operators is an example of a  $C^*$ -algebra with unity. An algebra is a vector space where multiplication is also defined (the product of two operators is an operator), and roughly speaking the  $*$  means that we can form Hermitian adjoints. The operator algebra is said to be generated by the  $c_p$  operators and the unit operator.

**Definition 4.2** (Contraction). A contraction between two arbitrary creation and annihilation operators  $X$  and  $Y$  is the number defined by

$$\overline{XY} \equiv \langle -|XY|-\rangle. \quad (4.21)$$

Alternatively,

$$\overline{XY} \equiv XY - \{XY\}. \quad (4.22)$$

The two definitions Eq. (4.21) and Eq. (4.22) are equivalent, see Exercise ??.

Let us list all the possible contractions:

$$\overline{c_p^\dagger c_q^\dagger} = \langle -|c_p^\dagger c_q^\dagger|-\rangle = 0 \quad (4.23a)$$

$$\overline{c_p c_q} = \langle -|c_p c_q|-\rangle = 0 \quad (4.23b)$$

$$\overline{c_p^\dagger c_q} = \langle -|c_p^\dagger c_q|-\rangle = 0 \quad (4.23c)$$

$$\overline{c_p c_q^\dagger} = \langle -|c_p c_q^\dagger|-\rangle = \delta_{pq}. \quad (4.23d)$$

As we see, most contractions are actually zero, and the only nonzero contractions are between annihilation operators *to the left* and creation operators *to the right*.

We also define contractions between two operators *inside a normal ordered product*.

**Definition 4.3** (Contractions inside normal-ordering operator). Let  $A_1 \cdots A_n$  be an operator product string, and let  $(x, y)$  be a pair of operators,  $1 \leq x < y \leq n$ . Let  $\sigma \in S_n$  be any permutation such that  $\sigma(1) = x$  and  $\sigma(2) = y$ . We define

$$\{A_1 \cdots \overline{A_x \cdots A_y} \cdots A_n\} \equiv (-1)^{|\sigma|} \{\overline{A_x A_y} A_{\sigma(3)} \cdots A_{\sigma(n)}\} \quad (4.24)$$

More generally, if we have  $m$  contracted pairs  $(x_i, y_i)$ ,  $x_i < y_i$ , we draw contraction lines (that may cross) between the pairs, and define the result as follows: Select a permutation  $\sigma \in S_n$  such that  $\sigma(1) = x_1$ ,  $\sigma(2) = y_1$ ,  $\sigma(3) = x_2$ ,  $\sigma(4) = y_2$ , etc., and set

$$\overbrace{\{A_1 A_2 \cdots A_n\}}^{m \text{ contraction lines}} = (-1)^{|\sigma|} \{\overline{A_{x_1} A_{y_1}} \cdots \overline{A_{x_m} A_{y_m}} A_{\sigma(2m+1)} \cdots A_{\sigma(n)}\}. \quad (4.25)$$

The permutation can be constructed by anticommuting each operator pair  $(x_i, y_i)$  to the front, collecting the number  $f_i$  of anticommutations. The total permutation sign is then  $(-1)^{f_1 + f_2 + \cdots + f_m}$ . The pairs may be treated in any order, the end result is the same.

Example, with one contracted pair:

$$\{A_1 \cdots \overline{A_x \cdots A_y} \cdots A_n\} \equiv (-1)^{x-1+y-2} \{\overline{A_x A_y} A_1 \cdots \cancel{A_x} \cdots \cancel{A_y} \cdots A_n\} = (-1)^{x+y+1} \overline{A_x A_y} \{A_1 \cdots \cancel{A_x} \cdots \cancel{A_y} \cdots A_n\} \quad (4.26)$$

Concrete examples (where each index  $i$  is short for  $p_i$ ):

$$\{c_1 c_2^\dagger \overline{c_3 c_4^\dagger}\} = \overline{c_3 c_4^\dagger} \{c_1 c_2^\dagger\} = -\delta_{34} c_1^\dagger c_2 \quad (4.27a)$$

$$\{\overline{c_1 c_2^\dagger} c_3 c_4^\dagger\} = \overline{c_1 c_2^\dagger} \{c_3 c_4^\dagger\} = -\delta_{12} c_3^\dagger c_4 \quad (4.27b)$$

$$\{\overline{c_1 c_2^\dagger c_3 c_4^\dagger}\} = \overline{c_1 c_4^\dagger} \{c_2^\dagger c_3\} = \delta_{14} c_2^\dagger c_3 \quad (4.27c)$$

$$\{c_1 c_2^\dagger c_3 c_4^\dagger\} = -c_2^\dagger c_4^\dagger \{c_1 c_3\} = 0 \quad (4.27d)$$

In the last example (4.27d), we could tell immediately that the result is zero, since any contraction between two creation operators vanishes.

Example with two contraction lines:

$$\{\overline{c_1 c_2^\dagger c_3 c_4^\dagger c_5 c_6^\dagger}\} = (-1)^2 \{\overline{c_1 c_4^\dagger} c_2^\dagger \overline{c_3 c_5 c_6^\dagger}\} = (-1)^3 \{\overline{c_1 c_4^\dagger} \overline{c_3 c_6^\dagger} c_2^\dagger c_5\} = -\delta_{14} \delta_{36} \{c_2^\dagger c_5\}. \quad (4.28)$$

Note how the contraction lines cross on the left-hand side. This is perfectly acceptable.

#### 4.1.4 Statement of Wick's Theorem

Wick's theorem states that every string of creation and annihilation operators can be written as a sum of normal-ordered products every possible contraction.

**Theorem 4.1** (Wick's Theorem). *Let  $A_1 \cdots A_n$  be an operator string of creation and annihilation operators. Then,*

$$\begin{aligned} A_1 A_2 \cdots A_n &= \{A_1 A_2 \cdots A_n\} + \sum_{(1)} \{A_1 \overline{\cdots} \cdots A_n\} + \sum_{(2)} \{A_1 \overline{\cdots} \cdots A_n\} \\ &+ \cdots + \sum_{(\lfloor \frac{n}{2} \rfloor)} \{A_1 \overbrace{\cdots \cdots \cdots A_n}^{\substack{\text{all contracted} \\ [n/2] \text{ contractions}}}\} \end{aligned} \quad (4.29)$$

The notation  $\sum_{(m)}$  signifies that we sum over all combinations of  $m$  contractions.

When  $n$  is even, the last sum signifies that we sum over  $n/2$  contractions, i.e., all operators are contracted. If  $n$  is odd, there is one uncontracted operator left in each term of the last sum.

#### 4.1.5 Vacuum expectation values using Wick's Theorem

Before we start with the proof of Wick's Theorem, we apply it to the evaluation of vacuum expectation values. For any string with at least one factor,

$$\langle - | \{A_1 \cdots A_n\} | - \rangle = 0. \quad (4.30)$$

This is so, because in the normal-order product, the annihilation operators are to the right, and the creation operators are on the left. For odd  $n$ , therefore, Wick's Theorem gives

$$\langle - | A_1 \cdots A_n | - \rangle = 0 \quad (n \text{ odd number}), \quad (4.31)$$

For even  $n$ ,

$$\langle - | A_1 A_2 \cdots A_n | - \rangle = \sum_{(\lfloor \frac{n}{2} \rfloor)} \{A_1 \overbrace{\cdots \cdots \cdots A_n}^{\substack{\text{all contracted}}}\}. \quad (4.32)$$

Since the only non-vanishing contractions are  $\overline{c_p c_q^\dagger} = \delta_{pq}$ , the number of contractions we need to consider when evaluating the sum is reduced. Moreover, if  $A_1 \cdots A_n$  contains a different number of creation and annihilation operators, at least one contraction of the form  $\overline{c_P c_Q}$  or  $\overline{c_P^\dagger c_Q^\dagger}$  must be present in Eq. (4.33), in every term, giving a zero expectation value at once.

Finally, one can show that the *sign* of a fully contracted operator product is  $(-1)^k$ , where  $k$  is the number of contraction line crossings. We will not prove this.

Clearly, Wick's theorem provides us with an algebraic method for rather easy determination of the terms that contribute to the matrix element.

We conclude with a recipe:

**Theorem 4.2** (Vacuum expectation values using Wick's Theorem). *Let  $A_1 \cdots A_n$  be a string of creation and annihilation operators.*

*If  $n$  is odd  $\langle -|A_1 \cdots A_n| - \rangle = 0$ .*

*Assume  $n$  is even. If  $A_1 \cdots A_n$  contains a different number of creation operators compared to annihilation operators,  $\langle -|A_1 \cdots A_n| - \rangle = 0$ .*

*Finally,*

$$\langle -|A_1 \cdots A_n| - \rangle = \sum_{\text{all contr.}} \{ \overline{A_1 A_2} \overline{A_3 A_4} \cdots \overline{A_k A_{k+1}} \overline{A_{k+2} A_{k+3}} \cdots A_n \}, \quad (4.33)$$

where the sum runs over all possible combinations of  $n/2$  contractions on the form

$$\overline{c_p c_q^\dagger}$$

The sign of each term in the sum is  $(-1)^k$ , where  $k$  is the number of crossings of contraction lines.

#### 4.1.6 Exercises

**Exercise 4.1.** Show that if a string  $A_1 \cdots A_n$  contains  $n_1$  creation operators and  $n_2$  annihilation operators in total, then there are  $n_1! n_2!$  strings that are on normal-ordered form, but that all the strings are the same as operators. △

**Exercise 4.2.** Prove that, for any permutation  $\sigma \in S_n$ ,

$$\{A_1 A_2 \cdots A_n\} = (-1)^{|\sigma|} \{A_{\sigma(1)} A_{\sigma(2)} \cdots A_{\sigma(n)}\}. \quad (4.34)$$

△

**Exercise 4.3.** Show Eq. (4.22) from Eq. (4.21), by considering the 4 possible cases. △

**Exercise 4.4.** Evaluate the following contractions inside normal-ordering operators:

1.  $\{ \overline{c_p^\dagger c_q} \overline{c_r c_s^\dagger} \}$
2.  $\{ \overline{c_{p_1} c_{p_2} \cdots c_{p_n} c_{q_1}^\dagger c_{q_2}^\dagger \cdots c_{q_m}^\dagger} \}$
3.  $\{ \overline{c_{12} c_{13} c_{01} c_1^\dagger c_2^\dagger c_3 c_4 c_5^\dagger c_6^\dagger c_7 c_8 c_9 c_{10}^\dagger c_{11}^\dagger} \}$

△

## 4.2 Proof of Wick's Theorem

The proof of Wick's theorem is by induction on the length  $n$  of the operator string. In mathematical induction, we prove a statement  $\mathcal{P}_n$  for all integers  $n$  by first proving it for  $n = 1$ , and then prove that  $\mathcal{P}_{n+1}$  must hold under the assumption that  $\mathcal{P}_n$  holds.

Here, the statement  $\mathcal{P}_n$  is (4.29).  $\mathcal{P}_1$  and  $\mathcal{P}_2$  are easily shown to be true (prove it!).

### 4.2.1 A lemma

We first prove a lemma.

**Lemma 4.1.** *Let  $A_r$ ,  $r = 1, \dots, n$  be creation and annihilation operators. Let  $B$  be a creation or annihilation operator. Then,*

$$\{A_1 A_2 \cdots A_n\} B = \sum_{r=1}^n \{A_1 A_2 \cdots \overline{A_r \cdots A_n B}\} + \{A_1 \cdots A_n B\}. \quad (4.35)$$

*Proof.* We have two cases: Assume first that  $B$  is an annihilation operator. Then all the contractions on the right-hand side vanish. Also,  $\{A_1 \cdots A_n\} B = \{A_1 \cdots A_n B\}$ .

Next case: Assume that  $B = c_q^\dagger$ , a creation operator.

We have again two cases to consider: first, that all the  $A_i = c_{p_i}$  are annihilation operators. In that case, we can verify that the left- and right-hand sides are equal by using the fundamental anticommutator relations. The left-hand side is equal to  $A_1 A_2 \cdots A_n B$ , since  $A_1 \cdots A_n$  is already a normal-ordered product. We use the definition of the contraction to pull  $B$  to the left of the expression to achieve normal order, at the cost of some extra terms:

$$\{A_1 A_2 \cdots A_n\} B = A_1 \cdots A_n B = A_1 \cdots A_{n-1} (\overline{A_n B} - B A_n), \quad (4.36)$$

since  $\{c_{p_i}, c_q^\dagger\} = \delta_{p_i, q} = \overline{c_{p_i} c_q^\dagger}$ . Continuing the anticommutation, pulling  $B$  to the left, we get

$$\begin{aligned} \{A_1 A_2 \cdots A_n\} B &= A_1 \cdots A_{n-1} \overline{A_n B} - A_1 \cdots A_{n-2} \overline{A_{n-1} B A_n} + A_1 \cdots A_{n-3} \overline{A_{n-2} B A_{n-1} A_n} - \cdots \\ &\quad + (-1)^n B A_1 \cdots A_n \\ &= \{A_1 \cdots A_{n-1} \overline{A_n B}\} + \{A_1 \cdots A_{n-2} \overline{A_{n-1} A_n B}\} + \{A_1 \cdots A_{n-3} \overline{A_{n-2} A_{n-1} A_n B}\} + \cdots \\ &\quad + (-1)^n B A_1 \cdots A_n \end{aligned} \quad (4.37)$$

This is equal to the right-hand side. This proves the case for all  $A_i$  annihilation operators, and it remains to prove it when we have creation operators in the mix.

Multiply Eq. (4.35) from the *left* by a creation operator  $A_0$ . We observe that normal order is preserved on the left hand side since  $A_0$  is a creation operator and can stand to the far right,

$$A_0 \{A_1 \cdots A_n\} = \{A_0 \cdots A_n\},$$

and similarly  $A_0 \{A_1 \cdots A_n\} B = \{A_0 \cdots A_n\} B$ . Also,

$$A_0 \sum_{r=1}^n \sum_{s=1}^n \{A_1 A_2 \cdots \overline{A_r \cdots A_n B}\} = \sum_{r=0}^n \{A_0 A_1 A_2 \cdots \overline{A_r \cdots A_n B}\},$$

since  $\overline{A_0 B} = 0$ . Thus, the statement of the lemma is true also when  $A_0$  is a creation operator. Clearly, we can continue, and add as many creation operators we like. Thus, the lemma is true for strings of the form  $C_1 \cdots C_k A_{k+1} \cdots A_n$ , where  $C_i$  are creation operators, and  $A_i$  are annihilation operators. By permuting this string, we gain a sign change on all terms, and the terms in the sum over  $r$  are reordered, but leaving the sum invariant. Thus, the lemma is proved for arbitrary strings  $A_1 \cdots A_n$ .  $\square$

We introduce another lemma, which generalizes Lemma 4.1 to the case where we have an arbitrary number  $m$  contractions between the  $n$  operators inside the normal order operator.

**Lemma 4.2.** *Suppose  $A_1 \cdots A_n$  is a given operator string, and suppose we choose  $m$  pairs  $(x_i, y_i)$  to contract from this string, with  $x_i < y_i$ . Let  $S = \{1, 2, \dots, N\} \setminus \{x_1, y_1, \dots, x_m, y_m\}$  be the remaining indices when all pairs are removed. Let  $B$  a creation or annihilation operator. Then,*

$$\{A_1 A_2 \cdots A_{n-1} A_n\} B = \{A_1 A_2 \cdots A_{n-1} A_n B\} + \sum_{r \in S} \{A_1 A_2 \cdots A_r \cdots A_{n-1} A_n B\} \quad (4.38)$$

where the notation indicates that all  $m$  pairs are contracted from the  $A_i$ s.

*Proof.* let  $S = \{1, 2, \dots, N\}$  (a set). The pairs are distinct, which we write mathematically as  $p_i \subset S \setminus (\cup_{j=1}^{i-1} p_j)$ .

Consider the left-hand side of Eq. (4.38). We perform the anticommutations that brings first the operators  $A_{x_1}$  and  $A_{y_1}$  to the front, then the next pair, etc. The first pair gives a sign  $(-1)^{f_1}$ , for  $f_1$  “flips”. The next pair gives a sign  $(-1)^{f_2}$ , and so on. Let  $C_1 \cdots C_{n-2m}$  be the remaining operators  $A_i$  with  $i \in S$ . We arrive at

$$\{A_1 A_2 \cdots A_{n-1} A_n\} = (-1)^{f_1 + f_2 + \dots + f_m} \overline{A_{x_1} A_{y_1}} \cdots \overline{A_{x_m} A_{y_m}} \{C_1 \cdots C_{n-2m}\}. \quad (4.39)$$

Now, by the first lemma,

$$\begin{aligned} \{A_1 A_2 \cdots A_{n-1} A_n\} B &= (-1)^{f_1 + f_2 + \dots + f_m} \overline{A_{x_1} A_{y_1}} \cdots \overline{A_{x_m} A_{y_m}} \left[ \{C_1 \cdots C_{n-2m} B\} \right. \\ &\quad \left. + \sum_{r=1}^{n-2m} \{C_1 \cdots C_r \cdots C_{n-2m} B\} \right] \end{aligned} \quad (4.40)$$

Consider the first term inside the bracket. We can move the contractions inside again, each pair passing the same operators as when extracted, but in reverse order, giving an overall sign change that cancels  $(-1)^{f_1 + \dots + f_m}$ , since no operator passes  $B$ . This reproduces the first term on the left-hand side of Eq. (4.38).

The same is actually true for the second term. This is seen by noting that, when having multiple contractions to resolve, as in the final term in Eq. (4.38), the order of the “extractions” of the pairs does not matter. The final one happens to be  $\overline{C_r B}$ , which is left inside.

This completes the proof.  $\square$

### 4.2.2 Proof of Wick's Theorem using the lemma

We now prove Wick's Theorem. Assume now that  $\mathcal{P}_n$  is true. Multiply Eq. (4.29) from the right by an operator  $A_{n+1}$ :

$$\begin{aligned}
A_1 A_2 \cdots A_n A_{n+1} &= \{A_1 A_2 \cdots A_n\} A_{n+1} + \sum_{(1)} \{\overline{A_1 A_2} \cdots A_n\} A_{n+1} \\
&+ \sum_{(2)} \{\overline{A_1 A_2 A_3 A_4} \cdots A_n\} A_{n+1} \\
&+ \cdots + \sum_{(\lfloor \frac{n}{2} \rfloor)} \{\overline{A_1 A_2 A_3 A_4} \cdots \overline{A_k A_{k+1} A_{k+2} A_{k+3}} \cdots A_n\} A_{n+1}
\end{aligned} \tag{4.41}$$

Each sum is a sum over  $m$  contractions, including the first where we have  $m = 0$ . We now use Lemma 4.2 and write

$$\begin{aligned}
\sum_{(m)} \{\overline{A_1 A_2 A_3 A_4} \cdots A_n\} A_{n+1} &= \sum_{(m)} \{\overline{A_1 A_2 A_3 A_4} \cdots A_n A_{n+1}\} + \sum_{(m)} \sum_r \{\overline{A_1 A_2 A_3 A_4} \cdots A_n A_{n+1}\} \\
&:= X_m + I_m.
\end{aligned} \tag{4.42}$$

Here, the sum over  $r$  is over the remaining indices after  $m$  pairs are contracted. The term  $X_m$  contains all possible  $m$  contractions *excluding*  $A_{n+1}$ , while  $I_m$  contains all possible  $m + 1$  contractions *including*  $A_{n+1}$ . We now get

$$A_1 \cdots A_{n+1} = X_0 + I_0 + X_1 + I_1 + \cdots X_{\lfloor n/2 \rfloor}. \tag{4.43}$$

Note that  $I_{\lfloor n/2 \rfloor} = 0$ , since there is no operator left to contract  $A_{n+1}$  with after  $2\lfloor n/2 \rfloor$  operators have been contracted.

Write

$$A_1 \cdots A_{n+1} = X_0 + (I_0 + X_1) + (I_1 + X_2) + \cdots X_{\lfloor n/2 \rfloor}. \tag{4.44}$$

and note that  $(I_m + X_{m+1})$  is the sum over *all possible*  $m + 1$  contractions of the string  $A_1 \cdots A_{n+1}$ . Thus, Wick's Theorem is proved.



### 4.3 Using Wick's Theorem

In this section, we see some examples of how to use Wick's Theorem to compute vacuum expectation values. First, we state, but do not prove, a theorem regarding the *sign* of a vacuum expectation value of a fully contracted normal-ordered product. The theorem simplifies enormously the work involved in computing the sign of the permutation needed to bring all the contracted pairs to the front.

**Theorem 4.3** (Sign rule for vacuum-expectation values). *Let  $A_1 \cdots A_n$  be an operator string of creation and annihilation operators, where  $n$  is even. Let  $n/2$  contractions be assigned, contracting  $A_{x_i}$  with  $A_{y_i}$  for all  $n/2$  pairs of operators,  $x_i < y_i$ , i.e., we have  $n/2$  contractions of the form  $\overline{A_{x_i} A_{y_i}}$ . Then,*

$$\langle - | A_1 A_2 A_3 \cdots A_{n-1} A_n | - \rangle = \overline{A_{x_1} A_{y_1}} \cdots \overline{A_{x_{n/2}} A_{y_{n/2}}} (-1)^s, \quad (4.45)$$

where  $s$  is the number of contraction line crossings on the left-hand side.

Let us compute a few vacuum expectation values with the aid of this rule, and also the simplifications we gain when we know that *all annihilation operators must be contracted with a creation operator to the right*.

We now also simplify the notation a bit, and write, in place of the ordinary creation and annihilation operators,

$$\mu^\dagger \equiv c_\mu^\dagger, \quad \mu = c_\mu.$$

Example 1:

$$\langle \mu_1 \cdots \mu_3 | \alpha^\dagger \beta | \mu_1 \cdots \mu_3 \rangle = \langle - | \mu_3 \mu_2 \mu_1 \alpha^\dagger \beta \mu_1^\dagger \mu_2^\dagger \mu_3^\dagger | - \rangle. \quad (4.46)$$

Example 2:

$$\langle \mu_1 \cdots \mu_3 | \alpha_1^\dagger \alpha_2^\dagger \beta_2 \beta_1 | \mu_1 \cdots \mu_3 \rangle = \langle - | \mu_3 \mu_2 \mu_1 \alpha_1^\dagger \alpha_2^\dagger \beta_2 \beta_1 \mu_1^\dagger \mu_2^\dagger \mu_3^\dagger | - \rangle. \quad (4.47)$$

(Finish the examples . . .)

#### 4.3.1 Exercises

**Exercise 4.5.** (Requires a fair amount of experience with permutations.) Prove the sign rule for the fully contracted terms.  $\triangle$

**Exercise 4.6.** Write out the statement of Wick's Theorem for the following operator strings, and simplify where you can:

1.  $c_p c_q^\dagger$
2.  $c_p^\dagger c_q c_r^\dagger c_s$
3.  $c_s c_t^\dagger c_u^\dagger c_p c_q c_r^\dagger$

$\triangle$

## 4.4 The Generalized Wick's Theorem

We here present a very useful generalization of Wick's Theorem. Even though Wick's Theorem greatly simplifies the evaluation of vacuum expectation values, it is a fact that most such expectation values one wants to compute are of strings of where substrings are already on normal-ordered form,

$$\{A_1 A_2 \cdots A_n\} \{B_1 B_2 \cdots B_m\} \cdots \{Z_1 Z_2 \cdots Z_x\} \quad (4.48)$$

For the standard Wick's Theorem case, all the substrings are of length 1.

Recall, that in the usual Wick's Theorem, we sum over normal-order products with contractions. This is also the case for the Generalized Wick's Theorem, but now each contraction must *involve operators from different substrings*. That is, contractions involving two operators from the same substring do not contribute.

This is in fact almost obvious: If a string is on normal order, all the contractions between the operators are zero, since annihilation operators are to the right of the creation operators.

**Theorem 4.4** (The Generalized Wick's Theorem). *Let  $A_1 \cdots A_n$ ,  $B_1 \cdots B_m$ , etc, be operator strings. Then,*

$$\begin{aligned} \{A_1 A_2 \cdots A_n\} \{B_1 B_2 \cdots B_m\} \cdots \{Z_1 Z_2 \cdots Z_x\} &= \{A_1 \cdots A_n : B_1 \cdots B_m : \cdots : Z_1 \cdots Z_x\} \\ &+ \sum'_{(1)} \{A_1 \cdots A_n : B_1 \cdots B_m : \cdots : Z_1 \cdots Z_x\} + \cdots + \sum'_{(\lfloor \frac{n}{2} \rfloor)} \underbrace{\{A_1 \cdots \cdots \cdots Z_x\}}_{\lfloor n/2 \rfloor \text{ contractions}} \end{aligned} \quad (4.49)$$

The notation  $\sum'_{(m)}$  signifies that we sum over all combinations of  $m$  contractions that each involve operators from different substrings. The vertical dots are here only a helpful device to remind ourselves which operators belong to which substrings.

When  $n$  is even, the last sum signifies that we sum over  $n/2$  contractions, i.e., all operators are contracted. The restriction to inter-string contractions implies that the maximum number of contractions in a term usually is smaller than  $\lfloor n/2 \rfloor$ .

Here is an example:

$$\begin{aligned} \{\hat{A}_1 \hat{A}_2 \hat{A}_3\} \{\hat{A}_4 \hat{A}_5\} &= \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} \\ &+ \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} \\ &+ \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} + \{\hat{A}_1 \hat{A}_2 \hat{A}_3 : \hat{A}_4 \hat{A}_5\} \end{aligned} \quad (4.50)$$

The vertical dots are for clarification only.

### 4.4.1 Exercises

**Exercise 4.7.** In this exercise,  $c_1^\dagger$  means  $c_{p_1}^\dagger$ , etc., for example  $\delta_{12} = \delta_{p_1 p_2}$ .

Write out, using the Generalized Wick's Theorem, the following operators. Use normal-ordering relative to the zero-particle vacuum  $|- \rangle$ . Start by identifying substrings already in normal-order product form.

Compare with the application of the original Wick's Theorem.

a)  $c_1 c_2 c_3^\dagger c_4^\dagger$

b)  $c_1 c_2^\dagger c_3^\dagger c_4 c_5 c_6^\dagger$

c)  $c_1 c_2 c_4^\dagger c_5 c_6 c_7^\dagger c_8 c_9$

$\Delta$

## 4.5 Normal ordering relative to Hartree–Fock vacuum: particle-hole formalism

In this section, we will introduce the concept of quasiparticles, or particle-hole formalism. Often, a single Slater determinant can be a good approximation, for example the Hartree–Fock Slater determinant. If this approximation is not good enough, one adds a correction on top of that. Therefore it makes sense to develop a convenient way to describe this small correction. Moreover, evaluating matrix elements of operators in  $N$ -particle Slater determinant bases is cumbersome with Wick’s theorem; the number of contractions grows quickly, increasing the opportunity for mistakes.

We start with dividing the single-particle functions  $\{\phi_p\}$  into two sets, the *occupied* and *unoccupied* (sometimes *virtual*) single-particle states. It is convenient to reserve the indices  $i, j, k, \dots$  for occupied, and  $a, b, c, \dots$  for unoccupied single-particle states.

Our *reference determinant*  $|\Phi\rangle$  is, by definition, constructed from precisely the occupied states,

$$|\Phi\rangle = \prod_{i \in \text{occ}} c_i^\dagger |-\rangle. \quad (4.51)$$

The complete Slater determinant basis can be constructed from the reference  $|\Phi\rangle$  by replacing one or more of the occupied states by unoccupied states. This is referred to as *excitations*. For example,

$$|\Phi_i^a\rangle \equiv c_a^\dagger c_i |\Phi\rangle \quad \text{single excitation}, \quad (4.52)$$

$$|\Phi_{ij}^{ab}\rangle \equiv c_b^\dagger c_j c_a^\dagger c_i |\Phi\rangle \quad \text{double excitation}, \quad (4.53)$$

et cetera. At most, we can replace  $N$  occupied indices.

We define *quasiparticle creation and annihilation operators* as follows:

$$b_i := c_i^\dagger, \quad b_a := c_a \quad (4.54)$$

with Hermitian adjoints

$$b_i^\dagger := c_i, \quad b_a^\dagger := c_a^\dagger. \quad (4.55)$$

Thus, for unoccupied indices, quasiparticle creation operators are the ordinary creation operators, but for occupied states, creating a quasiparticle is the same as destroying a particle in an occupied state. One says that  $b_i^\dagger$  creates a *hole*, while  $b_a^\dagger$  creates a *particle*.

It is an easy exercise to show that the anticommutator relations are preserved:

$$\{b_p, b_q^\dagger\} = \delta_{pq}, \quad \{b_p, b_q\} = 0. \quad (4.56)$$

Moreover, the reference  $|\Phi\rangle$  is the vacuum for the quasiparticle operators, i.e.,

$$b_p |\Phi\rangle = 0, \quad \text{for all } p. \quad (4.57)$$

We make the observation that the definition of normal-ordering, contractions, and ultimately Wick’s theorem, depended only on the anticommutator relations and the property  $c_p |-\rangle = 0$ . Thus, *Wick’s Theorem is valid also for quasiparticles*.

Using quasiparticle operators, excited determinants are written

$$|\Phi_i^a\rangle \equiv b_a^\dagger b_i^\dagger |\Phi\rangle \quad \text{single excitation}, \quad (4.58)$$

$$|\Phi_{ij}^{ab}\rangle \equiv b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger |\Phi\rangle \quad \text{double excitation}, \quad (4.59)$$

et cetera. The operator  $b_a^\dagger b_i^\dagger$  is often referred to as a *particle-hole pair creation operator*. Thus, an  $n$ -fold excited determinant contains  $n$  particle-hole pairs, i.e.,  $2n$  quasiparticles.

We see that applying Wick's Theorem to matrix elements involving at least low-order excited determinants should be *much simpler* than using the “bare” creation and annihilation operators.

Any wavefunction in with  $N$  particles can be written

$$|\Psi_N\rangle = C_0 |\Phi\rangle + \sum_{ia} C_i^a b_a^\dagger b_i |\Phi\rangle + \frac{1}{2!^2} \sum_{ijab} C_{ij}^{ab} b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger |\Phi\rangle + \dots, \quad (4.60)$$

where the factor  $1/2!^2$  comes from the double counting of the two particle-hole states. The sum extends all the way up to  $N$  particle hole pairs.

We can illustrate particles and holes with pictures. We let  $N = 4$ , and get the reference determinant

$$|\Phi\rangle = c_1^\dagger c_2^\dagger c_3^\dagger c_4^\dagger |-\rangle. \quad (4.61)$$

We can draw a picture like this:

$$|\Phi\rangle = \bullet \bullet \bullet \bullet \circ \circ \circ \circ \circ \circ$$

Let us create a quasiparticle:

$$b_i^\dagger |\Phi\rangle = c_i |123 \dots N\rangle = (-1)^{i-1} |123 \dots (i-1) (i+1) \dots N\rangle. \quad (4.62)$$

$$b_a^\dagger |\Phi\rangle = c_a^\dagger |123 \dots N\rangle = (-1)^N |123Na\rangle. \quad (4.63)$$

In pictures,

$$b_i^\dagger |\Phi\rangle = \bullet \bullet \circ_i \bullet \circ \circ \circ \circ \circ \circ$$

$$b_a^\dagger |\Phi\rangle = \bullet \bullet \bullet \bullet \circ \circ \circ \bullet_a \circ \circ$$

Note that  $b_i^\dagger |\Phi\rangle$  contains  $N - 1$  “real” particles, while  $b_a^\dagger |\Phi\rangle$  contains  $N + 1$  “real” particles. They contain, respectively, one hole and one particle.

Creating a particle-hole pair results in a state with  $N$  “real” particles, since  $b_a^\dagger b_i^\dagger = c_a^\dagger c_i$  preserves  $N$  when acting on a state. Acting on the reference, we get  $N - 1$  occupied single-particle functions below  $N$ , and 1 occupied single-particle function above  $N$ , in pictures,

$$b_a^\dagger b_i^\dagger |\Phi\rangle = \bullet \bullet \circ_i \bullet \circ \circ \circ \circ \bullet_a \circ \circ$$

Clearly, by creating another particle-hole pair with  $b_b^\dagger b_j^\dagger$ , we get a Slater determinant with two particles and two holes, in total  $N$  particles. We are left with  $N - 2$  “real” particles below  $N$ .

### 4.5.1 Exercices

**Exercise 4.8.** Prove the quasiparticle anticommutator relations. △

**Exercise 4.9.** Prove that

$$|\Phi_{ij}^{ab}\rangle = -|\Phi_{ij}^{ba}\rangle = -|\Phi_{ji}^{ab}\rangle = |\Phi_{ji}^{ba}\rangle. \quad (4.64)$$

More generally, let  $\sigma$  and  $\tau$  be two permutations of  $n \leq N$  indices. Prove that

$$|\Phi_{i_1 i_2 \dots i_n}^{a_1 a_2 \dots a_n}\rangle = (-1)^{|\sigma|} (-1)^{|\tau|} |\Phi_{i_{\sigma(1)} i_{\sigma(2)} \dots i_{\sigma(n)}}^{a_{\tau(1)} a_{\tau(2)} \dots a_{\tau(n)}}\rangle. \quad (4.65)$$

△

**Exercise 4.10.** Let there be in total  $L$  single-particle functions  $\phi_p$  given. If we restrict  $a \leq L$ , how many *linearly independent* two-particle-two-hole determinants can you create, with  $N$  particles in the reference? △

**Exercise 4.11.** Compute using Wick's theorem relative to the vacuum  $|- \rangle$  and relative to the vacuum  $|\Phi\rangle$ , and compare the method and amount of work.

$$\langle \Phi | c_p^\dagger c_q^\dagger c_s c_r | \Phi \rangle. \quad (4.66)$$

Note that you get several cases when using quasiparticles, depending on the indices  $pqsr$  being occupied or unoccupied. △

**Exercise 4.12.** Use Wick's Theorem with respect to quasiparticles and write down the following operators as a sum of normal-ordered strings with as few terms as possible (i.e., only include nonvanishing contractions):

- a)  $b_a b_i^\dagger$
- b)  $b_i^\dagger b_c b_a^\dagger$
- c)  $b_i b_j b_a^\dagger b_b^\dagger b_c b_k^\dagger$
- d)  $b_a b_i b_j b_b^\dagger b_c^\dagger b_d b_k^\dagger$

△

**Exercise 4.13.** Compute using the quasiparticle version of Wick's Theorem:

- a)  $\langle \Phi | \hat{H}_0 | \Phi \rangle$ , with  $\hat{H}_0 = \sum_{pq} h_q^p c_p^\dagger c_q$ .
- b)  $\langle \Phi_i^a | \hat{H}_0 | \Phi \rangle$
- c)  $\langle \Phi_i^a | \hat{H}_0 | \Phi_j^b \rangle$
- d)  $\langle \Phi_{ij}^{ab} | \hat{H}_0 | \Phi \rangle$

△

## 4.6 Operators on normal-order form

### 4.6.1 The number operator

We introduce the normal-ordered form of second-quantized operators such as  $\hat{H} = \hat{H}_0 + \hat{W}$ , relative to quasiparticle vacuum. I.e., we want the operator to be written such that all quasiparticle annihilation operators are to the right, and creation operators to the left. This is achieved using Wick's Theorem, and results in the original operator obtaining more terms. On the other hand, computations of matrix elements becomes more systematic, as we can employ the generalized Wick's Theorem.

We start with the number operator  $\hat{N}$ , as an easy warm-up. First, we rewrite the second-quantized operator using quasiparticle operators:

$$\hat{N} = \sum_p c_p^\dagger c_p = \sum_i c_i^\dagger c_i + \sum_a c_a^\dagger c_a = \sum_i b_i b_i^\dagger + \sum_a b_a^\dagger b_a. \quad (4.67)$$

We now use Wick's theorem, *relative to quasiparticle operators*, to get

$$\begin{aligned} \hat{N} &= \sum_i \left[ \{b_i b_i^\dagger\} + \overline{b_i b_i^\dagger} \right] + \sum_a \left[ \{b_a^\dagger b_a\} + \overline{b_a^\dagger b_a} \right] \\ &= \sum_i \left[ -b_i^\dagger b_i + 1 \right] + \sum_a b_a^\dagger b_a \\ &= N - \sum_i b_i^\dagger b_i + \sum_a b_a^\dagger b_a. \end{aligned} \quad (4.68)$$

This is the normal-ordered form of  $\hat{N}$ . Interpreting, the last equality counts  $N$  minus the number of holes plus the number of particles.

We note how  $\hat{N}$  is split into a zero-quasiparticle part  $N$  (a constant) and a one-quasiparticle part  $\hat{N} - N$ . This is a general feature when we normal-order with respect to a new vacuum.

Let us act with  $\hat{N}$  on the quasiparticle vacuum, and observe:

$$\hat{N} |\Phi\rangle = \left( N - \sum_i b_i^\dagger b_i + \sum_a b_a^\dagger b_a \right) |\Phi\rangle = N |\Phi\rangle. \quad (4.69)$$

All terms but the constant term vanish, since quasiparticle vacuum contains no quasiparticles. The normal-ordered form has destruction operators to the left.

### 4.6.2 One-body operators

We continue with an arbitrary one-body operator

$$\hat{H}_0 = \sum_{pq} h_{pq}^0 c_p^\dagger c_q. \quad (4.70)$$

Introducing the quasiparticle operators at this stage leads to four distinct contributions to the operator, corresponding to the different  $pq = ij, ia, ai$ , and  $ab$  subsums. However, it is more convenient to use Wick's Theorem on the above  $\hat{H}_0$  expression without changing the creation- and annihilation operator notation. Thus, beware, when we normal order now, it is *relative to quasiparticles*.

Wick's Theorem gives

$$c_p^\dagger c_q = \{c_p^\dagger c_q\} + \overline{c_p^\dagger c_q}. \quad (4.71)$$

For a nonzero contraction, we need both  $PQ$  to be hole indices. The only surviving contractions are  $pq = ii$ , that is,  $\overline{b_i} b_i^\dagger$ . The reader should verify that the rest of the possible contractions vanish identically. Thus,

$$\begin{aligned}\hat{H}_0 &= \sum_{pq} h_q^p \{c_p^\dagger c_q\} + \sum_{pq} h_q^p \overline{c_p^\dagger} c_q \\ &= \sum_{pq} h_q^p \{c_p^\dagger c_q\} + \sum_i h_i^i \\ &= \hat{H}_0^{(1qp)} + \hat{H}_0^{(0qp)}.\end{aligned}\tag{4.72}$$

Note that  $\hat{H}_0$  is separated into a one-quasiparticle part and a constant zero-quasiparticle part. Explicitly,

$$\hat{H}_0^{(0qp)} = \sum_i h_i^i,\tag{4.73}$$

and

$$\hat{H}_0^{(1qp)} = - \sum_{ij} h_j^i b_j^\dagger b_i + \sum_{AI} h_i^a b_a^\dagger b_i^\dagger + \sum_{ia} h_a^i b_i b_a + \sum_{ab} h_b^a b_a^\dagger b_b,\tag{4.74}$$

where we have expanded the sum over  $pq$ .

### 4.6.3 Two-body operators

We continue with an arbitrary two-body operator

$$\hat{W} = \frac{1}{4} \sum_{PQRS} w_{RS}^{PQ} c_P^\dagger c_Q^\dagger c_S c_R, \quad w_{RS}^{PQ} \equiv \langle PQ | \hat{W} | RS \rangle_{AS},\tag{4.75}$$

where we assume that  $w_{RS}^{PQ}$  is anti-symmetrized. (This is at odds with earlier notation, but we do this to save space in the current section.)

Wick's Theorem gives

$$\begin{aligned}c_P^\dagger c_Q^\dagger c_S c_R &= \{c_P^\dagger c_Q^\dagger c_S c_R\} + \{\overline{c_P^\dagger} c_Q^\dagger c_S c_R\} + \{\overline{c_P^\dagger} \overline{c_Q^\dagger} c_S c_R\} + \{\overline{c_P^\dagger} \overline{c_Q^\dagger} \overline{c_S c_R}\} \\ &\quad + \{c_P^\dagger c_Q^\dagger \overline{c_S c_R}\} + \{c_P^\dagger c_Q^\dagger \overline{c_S} c_R\} + \{c_P^\dagger c_Q^\dagger \overline{c_S} \overline{c_R}\} \\ &\quad + \{\overline{c_P^\dagger} c_Q^\dagger \overline{c_S c_R}\} + \{\overline{c_P^\dagger} c_Q^\dagger \overline{c_S} c_R\} + \{\overline{c_P^\dagger} c_Q^\dagger \overline{c_S} \overline{c_R}\} \\ &= \{c_P^\dagger c_Q^\dagger c_S c_R\} + \overline{c_P^\dagger} c_Q^\dagger \{c_S c_R\} - \overline{c_P^\dagger} c_S \{c_Q^\dagger c_R\} + \overline{c_P^\dagger} c_R \{c_Q^\dagger c_S\} \\ &\quad + \overline{c_Q^\dagger} c_S \{c_P^\dagger c_R\} - \overline{c_Q^\dagger} c_R \{c_P^\dagger c_S\} + \overline{c_S c_R} \{c_P^\dagger c_Q^\dagger\} \\ &\quad + \overline{c_P^\dagger} \overline{c_Q^\dagger} \overline{c_S c_R} - \overline{c_P^\dagger} \overline{c_S} \overline{c_Q^\dagger} c_R + \overline{c_Q^\dagger} \overline{c_S} \overline{c_P^\dagger} c_R\end{aligned}\tag{4.76}$$

We see immediately, that analogously to the one-body operator, we will get

$$\hat{W} = \hat{W}^{(2qp)} + \hat{W}^{(1qp)} + \hat{W}^{(0qp)}.\tag{4.77}$$

The two-quasiparticle term is

$$\hat{W}^{(2qp)} = \frac{1}{4} \sum_{PQRS} w_{RS}^{PQ} \{c_P^\dagger c_Q^\dagger c_S c_R\}.\tag{4.78}$$



For the one-quasiparticle term, the singly contracted terms are either vanishing (the first and the sixth), or have equal value (the second through fifth), adding up to

$$\hat{W}^{(1qp)} = \sum_{PQI} w_{QI}^{PI} \{c_P^\dagger c_Q\} \quad (4.79)$$

while the constant term is

$$\hat{W}^{(0qp)} = \frac{1}{2} \sum_{IJ} w_{IJ}^{IJ}. \quad (4.80)$$

We now make an observation regarding contractions of *standard* creation and annihilation operators relative to *quasiparticle vacuum*. First,

$$\overline{c_P c_Q} = 0, \quad (4.81)$$

since the only possibility of a nonzero contraction is when the left and right operators are quasiparticle annihilation and creation operators, respectively, implying that  $P$  is an unoccupied index and  $Q$  is an occupied index. Such indices are always different. Similarly,

$$\overline{c_P^\dagger c_Q^\dagger} = 0. \quad (4.82)$$

The only nonzero contractions are

$$\overline{c_I^\dagger c_J} = \delta_{IJ}, \quad (4.83)$$

and

$$\overline{c_{AC}^\dagger c_B} = \delta_{AB}. \quad (4.84)$$

This can be of some use in practical calculations.

#### 4.6.4 Normal-ordered two-body Hamiltonian

Consider the full Hamiltonian on the form

$$\hat{H} = \hat{H}_0 + \hat{W}. \quad (4.85)$$

The Hamiltonian is normal-ordered relative to “real” particles. In terms of quasiparticles, we saw in the previous sections that we could split

$$\hat{H} = \hat{H}_0^{(0qp)} + \hat{W}^{(0qp)} + \hat{H}_0^{(1qp)} + \hat{W}^{(1qp)} + \hat{W}^{(2qp)}, \quad (4.86)$$

separating  $\hat{H}$  into zero, one and two-quasiparticle contributions. These were normal-ordered relative to quasiparticles.

It is conventional to write

$$\hat{H} = \hat{H}_N + E_0 = \hat{H}_N + \hat{W}_N + E_0, \quad (4.87)$$

where now  $\hat{H}_N$  is a sum of strings on normal-order form,

$$\hat{H}_N = \{\hat{F}\} + \{\hat{W}\} \quad (4.88)$$

with

$$E_0 = \hat{H}_0^{(0qp)} + \hat{W}^{(0qp)} = \sum_I h_I^I + \frac{1}{2} \sum_{ij} w_{ij}^{IJ}, \quad (4.89a)$$

$$\{\hat{F}\} = \hat{H}_0^{(1qp)} + \hat{W}^{(1qp)} = \sum_{PQ} (h_Q^P + \sum_I w_{QI}^{PI}) \{c_P^\dagger c_Q\}, \quad (4.89b)$$

$$\{\hat{W}\} = \hat{W}^{(2qp)} = \frac{1}{4} \sum_{PQRS} w_{RS}^{PQ} \{c_P^\dagger c_Q^\dagger c_S c_R\}. \quad (4.89c)$$

Thus,  $\{\hat{F}\}$  is the *total* one-quasiparticle operator part of  $\hat{H}$ , and contains contributions from  $\hat{W}$  as well as  $\hat{H}_0$ , gathered in an operator  $\hat{F}$ , while  $\{\hat{W}\}$  is the total two-quasiparticle operator part of  $\hat{H}$ . It coincides with simply normal-ordering  $\hat{W}$ .

#### 4.6.5 Full expressions for the normal-ordered Hamiltonian

For completeness, we expand  $\hat{H}_{0,N}$  and  $\hat{W}_N$  in terms of quasiparticle operators. This gives a lot of terms, especially in the two-body case. We start with  $\hat{H}_{0,N}$ , splitting the sum over  $pq$  into four terms:

$$\begin{aligned}\{\hat{F}\} &= \sum_{ij} f_j^i \{c_i^\dagger c_j\} + \sum_{ai} f_i^a \{c_a^\dagger c_i\} + \sum_{ia} f_a^i \{c_i^\dagger c_a\} + \sum_{ab} f_b^a \{c_a^\dagger c_b\} \\ &= \sum_{ai} f_i^a b_a^\dagger b_i^\dagger + - \sum_{ij} f_j^i b_j^\dagger b_i + \sum_{ab} f_b^a b_a^\dagger b_b + \sum_{ia} f_a^i b_i b_a +\end{aligned}\quad (4.90)$$

where

$$f_q^p = h_q^p + \sum_j w_{qj}^{pj}. \quad (4.91)$$

Next, we resolve the two-body operator. There are 16 terms:

$$\begin{aligned}\{\hat{W}\} &= \frac{1}{4} \sum_{ijkl} w_{kl}^{ij} \{c_i^\dagger c_j^\dagger c_l c_k\} + \frac{1}{4} \sum_{ijka} w_{ka}^{ij} \{c_i^\dagger c_k^\dagger c_a c_k\} + \frac{1}{4} \sum_{ijak} w_{ak}^{ij} \{c_i^\dagger c_j^\dagger c_k c_a\} + \frac{1}{4} \sum_{iajk} w_{jk}^{ia} \{c_i^\dagger c_a^\dagger c_k c_j\} \\ &+ \frac{1}{4} \sum_{aijk} w_{jk}^{ai} \{c_a^\dagger c_i^\dagger c_k c_j\} + \frac{1}{4} \sum_{ijab} w_{ab}^{ij} \{c_i^\dagger c_j^\dagger c_b c_a\} + \frac{1}{4} \sum_{iajb} w_{jb}^{ia} \{c_i^\dagger c_a^\dagger c_b c_j\} + \frac{1}{4} \sum_{iabj} w_{bj}^{ia} \{c_i^\dagger c_a^\dagger c_j c_b\} \\ &+ \frac{1}{4} \sum_{aijb} w_{jb}^{ai} \{c_a^\dagger c_i^\dagger c_b c_j\} + \frac{1}{4} \sum_{aibj} w_{bj}^{ai} \{c_a^\dagger c_i^\dagger c_j c_b\} + \frac{1}{4} \sum_{abij} w_{ij}^{ab} \{c_a^\dagger c_b^\dagger c_j c_i\} + \frac{1}{4} \sum_{iabc} w_{bc}^{ia} \{c_i^\dagger c_a^\dagger c_c c_b\} \\ &+ \frac{1}{4} \sum_{aibc} w_{bc}^{ai} \{c_a^\dagger c_i^\dagger c_c c_b\} + \frac{1}{4} \sum_{abci} w_{ic}^{ab} \{c_a^\dagger c_b^\dagger c_c c_i\} + \frac{1}{4} \sum_{abci} w_{ci}^{ab} \{c_a^\dagger c_b^\dagger c_i c_c\} + \frac{1}{4} \sum_{abcd} w_{cd}^{ab} \{c_a^\dagger c_b^\dagger c_d c_c\}\end{aligned}\quad (4.92)$$

Some terms are equal, and we rearrange the expression to read:

$$\begin{aligned}\{\hat{W}\} &= \frac{1}{4} \sum_{abij} w_{ij}^{ab} b_a^\dagger b_b^\dagger b_j^\dagger b_i^\dagger + \frac{1}{2} \sum_{abci} w_{ci}^{ab} b_a^\dagger b_b^\dagger b_i^\dagger b_c + \frac{1}{4} \sum_{aibj} w_{bj}^{ai} b_a^\dagger b_j^\dagger b_b^\dagger b_i + \frac{1}{2} \sum_{aijk} w_{jk}^{ai} b_a^\dagger b_k^\dagger b_j^\dagger b_i \\ &+ \frac{1}{4} \sum_{ijkl} w_{kl}^{ij} b_l^\dagger b_k^\dagger b_i b_j - \frac{1}{4} \sum_{iajb} w_{jb}^{ia} b_a^\dagger b_j^\dagger b_i b_b + \frac{1}{2} \sum_{iabj} w_{bj}^{ia} b_a^\dagger b_j^\dagger b_i b_b + \frac{1}{4} \sum_{abcd} w_{cd}^{ab} b_a^\dagger b_b^\dagger b_d b_c \\ &+ \frac{1}{2} \sum_{ijak} w_{ak}^{ij} b_k^\dagger b_i b_j b_a + \frac{1}{2} \sum_{aibc} w_{bc}^{ai} b_a^\dagger b_i b_c b_b + \frac{1}{4} \sum_{ijab} w_{ab}^{ij} b_i b_j b_b b_a\end{aligned}\quad (4.93)$$

**Exercise 4.14.** Verify that Eq. (4.92) equals Eq. (4.93). △

**Exercise 4.15.** Verify that  $\hat{W}_N$  in Eq. (4.93) is Hermitian, given that  $\hat{W}$  is Hermitian. △

**Exercise 4.16.** Consider a three-body operator

$$\hat{X} = \frac{1}{36} \sum_{pqrstu} x_{stu}^{pqr} c_p^\dagger c_q^\dagger c_r^\dagger c_u c_t c_s, \quad (4.94)$$

where  $x_{stu}^{pqr} = \langle pqr | \hat{x}(1, 2, 3) | stu \rangle$  is permutation antisymmetric in the upper and lower indices separately, i.e., it is the matrix of the three-particle operator  $x$ . Such operators occur in nuclear physics.

Compute the separation

$$\hat{X} = \hat{X}^{(0qp)} + \hat{X}^{(1qp)} + \hat{X}^{(2qp)} + \hat{X}^{(3qp)}. \quad (4.95)$$

Given a Hamiltonian is given by  $\hat{H} = \hat{H}_0 + \hat{W} + \hat{X}$ , write down the normal-ordered Hamiltonian, and identify the zero-quasiparticle, one-quasiparticle, two-quasiparticle, and three-quasiparticle parts. Hint:  $\hat{X}$  will contribute to all these.

△

## Chapter 5

# Hartree–Fock theory part II

### 5.1 Normal-ordered Hamiltonian using HF basis

When we normal-order our operators using the HF determinant as reference, we obtain a nice representation of our Hamiltonian. Recall the normal-ordered form for the Hamiltonian:

$$\begin{aligned}
 \hat{H} &= \hat{H}_0 + \hat{W} = [\hat{H}_0^{(0qp)} + \hat{W}^{(0qp)}] + [\hat{H}_0^{(1qp)} + \hat{W}^{(1qpP)}] + \hat{W}^{(2qp)} \\
 &= \left( \sum_i h_i^i + \frac{1}{2} \sum_{ij} w_{ij}^{ij} \right) + \{\hat{H}_0\} + \sum_{pq} \left( \sum_i w_{qi}^{pi} \right) \{c_p^\dagger c_q\} + \{\hat{W}\} \\
 &= E_{\text{HF}} + \{\hat{H}_0 + \hat{V}^{\text{HF}}\} + \{\hat{W}\} \\
 &= E_{\text{HF}} + \{\hat{F}\} + \{\hat{W}\}.
 \end{aligned} \tag{5.1}$$

Here, we have used that  $\{\dots\}$  is defined for linear combinations, too,

$$\{\hat{F}\} = \sum_{pq} f_q^p \{c_p^\dagger c_q\}, \quad f_q^p = h_q^p + \sum_i w_{qi}^{pi}, \tag{5.2}$$

$$\{\hat{W}\} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \{c_p^\dagger c_q^\dagger c_s c_r\}. \tag{5.3}$$

It is customary to introduce the normal-ordered Hamiltonian  $\hat{H}_N$  as

$$\hat{H} = \hat{H}_N + E_{\text{HF}}, \quad \hat{H}_N = \{\hat{F}\} + \{\hat{W}\}. \tag{5.4}$$

The condition that  $|\Phi\rangle$  solves the HF equations imply that

$$\{\hat{F}\} = \sum_{ij} f_j^i \{c_i^\dagger c_j\} + \sum_{ab} f_b^a \{c_a^\dagger c_b\} = - \sum_{ij} f_j^i \{b_j^\dagger b_i\} + \sum_{ab} f_b^a \{b_a^\dagger b_b\}. \tag{5.5}$$

In the canonical representation we obtain a further simplification,

$$\{\hat{F}\} = - \sum_i \epsilon_i \{b_i^\dagger b_i\} + \sum_a \epsilon_a \{b_a^\dagger b_a\}. \tag{5.6}$$

## 5.2 Brillouin's Theorem

What is special about the HF reference, is of course that it is chosen to be optimize a certain aspect of the basis, namely that the reference state has minimal energy. This has a reformulation in terms of second-quantization, namely *Brillouin's Theorem*:

**Theorem 5.1** (Brillouin's Theorem). *Let an orthonormal single-particle basis  $\{\phi_p\}$  be given, defining a basis of Slater determinants. Then,*

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = 0, \quad \forall i, a \quad (5.7)$$

*if and only if the single-particle basis functions solve the (non-canonical) HF equations.*

*Proof.* The theorem is proved by establishing

$$\langle \Phi_i^a | \hat{H} | \Phi \rangle = f_i^a. \quad (5.8)$$

Now,  $f_i^a = 0$  if and only if the non-canonical HF equations are satisfied by the single-particle basis. To establish (5.8), we write  $\hat{H} = E_0 + \{\hat{F}\} + \{\hat{W}\}$ , and note that  $\langle \Phi_i^a | \{\hat{W}\} | \Phi \rangle = 0$ , since  $\{\hat{W}\}$  is a two-quasiparticle operator, and there is no way to contract all four operators. Moreover,  $\langle \Phi_i^a | E_{\text{HF}} | \Phi \rangle = 0$ , too. The remaining matrix element is

$$\langle \Phi_i^a | \{\hat{F}\} | \Phi \rangle = \sum_{pq} f_q^p \langle \Phi | \overline{\{b_i b_a\} \{c_p^\dagger c_q\}} | \Phi \rangle = f_i^a. \quad (5.9)$$

□

### 5.3 Stability of HF

The SCF procedure for solving the HF equations may or may not yield a local or global minimum for the energy. It may, however, also be a saddle point. In that case, the HF energy is *unstable*, as small perturbations in the HF single-particle functions may lower the energy further. In this section, we derive the condition that the HF energy is stable under all small perturbations.

From Exercise 2.22 we know that any Slater determinant can be obtained from  $|\Phi\rangle$  using an exponential operator,

$$|\Phi'\rangle = e^{\hat{X}} |\Phi\rangle, \quad (5.10)$$

where  $\hat{X} = \sum_{pq} x_{pq} c_p^\dagger c_q$ , and where  $\hat{X}^\dagger = -\hat{X}$ . The reason is that  $e^x$  is the unitary change-of-basis matrix from the basis  $\{\phi_p\}$  to the new basis  $\{\phi'_p\}$ ,

$$\phi'_p = \sum_q \phi_q (e^x)_{qp}. \quad (5.11)$$

Thus, may express perturbations of the HF energy as

$$\Delta E = \langle \Phi' | \hat{H} | \Phi' \rangle - E_{\text{HF}} = \langle \Phi | e^{-\hat{X}} \hat{H}_N e^{\hat{X}} | \Phi \rangle = \langle \Phi | [\hat{H}_N, \hat{X}] | \Phi \rangle + \frac{1}{2} \langle \Phi | [[\hat{H}_N, \hat{X}], \hat{X}] | \Phi \rangle + O(x^3). \quad (5.12)$$

Here, we used the Baker–Campbell–Hausdorff formula for the similarity transformation. The first-order term vanishes by Brillouin’s Theorem; see exercise 5.1. Indeed, that the linear term vanishes is nothing but the stationarity condition for the energy, and must vanish at the HF solution.

Next, we consider the second-order term.

$$\begin{aligned} \langle \Phi | [[\hat{H}_N, \hat{X}], \hat{X}] | \Phi \rangle &= \langle \Phi | \hat{H}_N \hat{X}^2 | \Phi \rangle + \langle \Phi | \hat{X}^2 \hat{H}_N | \Phi \rangle - 2 \langle \Phi | \hat{X} \hat{H}_N \hat{X} | \Phi \rangle \\ &\equiv A + A' - 2B. \end{aligned} \quad (5.13)$$

Each of the terms  $A$ ,  $A'$  and  $B$  have one- and two-body parts. To compute these, it is convenient to express  $\hat{X}$  on normal order,

$$\hat{X} = \sum_{pq} x_{pq} (\{c_p^\dagger c_q\} + \delta_{pq}^{\text{occ}}) = \sum_i x_{ii} + \{\hat{X}\}. \quad (5.14)$$

Here, we introduced the superscript “occ” on the Kronecker delta to indicate that in addition to  $p = q$  we must have  $p \in \text{occ}$ . (We similarly introduce  $\delta_{pq}^{\text{vir}}$ .) Indeed, as we compute only commutators, the constant term can be neglected, and we replace  $\hat{X}$  by  $\{\hat{X}\}$  from now on.

Let us consider first the term  $A$ . The one-body part vanishes. To see this, compute

$$A^{(1)} = \langle \Phi | \{\hat{F}\} \{\hat{X}\} \{\hat{X}\} | \Phi \rangle = \sum_{tu} f_u^t \langle \Phi | \{c_t^\dagger c_u\} \{\hat{X}\} \{\hat{X}\} | \Phi \rangle, \quad (5.15)$$

and note that  $c_t^\dagger$  must be contracted to the right, forcing  $t$  to be an occupied index. Similarly,  $u$  must be a virtual index. Thus, we obtain only matrix elements  $f_a^i$ , but these are zero due to  $|\Phi\rangle$  being the HF determinant. Similarly,  $(A')^{(1)} = 0$ .

$$\begin{aligned} A^{(2)} &= \sum_{pq} \sum_{rs} x_{pq} x_{rs} \langle \Phi | \{\hat{W}\} \{c_p^\dagger c_q\} \{c_r^\dagger c_s\} | \Phi \rangle \\ &= \frac{1}{4} \sum_{pq} \sum_{rs} \sum_{tuvw} x_{pq} x_{rs} w_{vw}^{tu} \langle \Phi | \{c_t^\dagger c_u^\dagger c_w c_v\} \{c_p^\dagger c_q\} \{c_r^\dagger c_s\} | \Phi \rangle \end{aligned} \quad (5.16)$$

There are four possible contractions.  $c_w$  and  $c_v$  must be contracted with  $c_p^\dagger$  and  $c_r^\dagger$ , and  $c_t^\dagger$  and  $c_u^\dagger$  must be contracted with  $c_q$  and  $c_s$ . One of these four contractions are

$$\langle \Phi | \{ \overbrace{c_t^\dagger c_u^\dagger c_w c_v}^{\text{contraction}} \} \{ c_p^\dagger c_q \} \{ c_r^\dagger c_s \} | \Phi \rangle = -\delta_{ts}^{\text{occ}} \delta_{uq}^{\text{occ}} \delta_{wr}^{\text{vir}} \delta_{vp}^{\text{vir}}. \quad (5.17)$$

Computing the three other contractions and inserting gives 4 identical terms when we use permutational antisymmetry of  $w_{vu}^{tu}$ , giving

$$A^{(2)} = \sum_{ai} \sum_{bj} x_{ai} x_{bj} w_{ab}^{ij}. \quad (5.18)$$

Similarly,  $(A')^{(2)}$  becomes

$$(A')^{(2)} = \sum_{ai} \sum_{bj} x_{ia} x_{jb} w_{ij}^{ab}. \quad (5.19)$$

The reader is encouraged to fill in the details on this calculation, see Exercise 5.2.

Next up is the term  $B = B^{(1)} + B^{(2)}$ . We compute

$$\begin{aligned} B^{(1)} &= \langle \Phi | \{ \hat{X} \} \{ \hat{F} \} \{ \hat{X} \} | \Phi \rangle = \sum_{pq} \sum_{rs} \sum_{tu} x_{pq} x_{rs} f_u^t \langle \Phi | \{ c_p^\dagger c_q \} \{ c_t^\dagger c_u \} \{ c_r^\dagger c_s \} | \Phi \rangle \\ &= \sum_{pq} \sum_{rs} \sum_{tu} x_{pq} x_{rs} f_u^t (\delta_{qt}^{\text{vir}} \delta_{ur}^{\text{vir}} \delta_{ps}^{\text{occ}} - \delta_{pu}^{\text{occ}} \delta_{ts}^{\text{occ}} \delta_{qr}^{\text{vir}}) \\ &= \sum_{ia} \sum_{jb} x_{ia} x_{bj} (\delta_{ij} f_b^a - \delta_{ab} f_i^j) \end{aligned} \quad (5.20)$$

The reader is encouraged to check the contractions. Finally,

$$\begin{aligned} B^{(2)} &= \langle \Phi | \{ \hat{X} \} \{ \hat{W} \} \{ \hat{X} \} | \Phi \rangle = \frac{1}{4} \sum_{pq} \sum_{rs} \sum_{tuvw} x_{pq} x_{rs} w_{vw}^{tu} \langle \Phi | \{ c_p^\dagger c_q \} \{ c_t^\dagger c_u^\dagger c_w c_v \} \{ c_r^\dagger c_s \} | \Phi \rangle \\ &= - \sum_{pq} \sum_{rs} \sum_{tuvw} x_{pq} x_{rs} w_{vw}^{tu} \delta_{qt}^{\text{vir}} \delta_{vr}^{\text{vir}} \delta_{pw}^{\text{occ}} \delta_{us}^{\text{occ}} = - \sum_{ia} \sum_{jb} x_{ia} x_{bj} w_{bi}^{aj}, \end{aligned} \quad (5.21)$$

where again we have in total 4 identical contributions. It is recommended that the reader does exercise 5.2, filling out the details.

We collect the terms, and find, up to second order in  $\hat{X}$ ,

$$\Delta E_{\text{HF}} = \frac{1}{2} \sum_{ia} \sum_{jb} \left[ x_{ai} x_{bj} w_{ab}^{ij} + x_{ia} x_{jb} w_{ij}^{ab} - 2x_{ia} x_{bj} (\delta_{ij} f_b^a - \delta_{ab} f_i^j - w_{bi}^{aj}) \right]. \quad (5.22)$$

Notably, only the off-diagonal blocks  $x_{ia}$  and  $x_{ai}$  enter this equation –  $x_{ij}$  and  $x_{ab}$  have no bearing on the result.

Since  $\hat{X}^\dagger = -\hat{X}$ , we have  $x_{ia} = -x_{ai}^*$ . Write  $x_\mu = x_{ia}$ , where  $\mu = (ia)$  is now a compound index. Let  $K_{\mu\nu} = K_{\nu\mu} = w_{ij}^{ab}$ ,  $L_{\mu\nu} = L_{\nu\mu}^* = -w_{bi}^{aj}$ , and  $D_{\mu\nu} = \delta_{ij} f_b^a - \delta_{ab} f_i^j$ . Notably,  $K = K^T$ ,  $L = L^H$ , and  $D = D^H$  as matrices. We obtain

$$\begin{aligned} \Delta E_{\text{HF}} &= \frac{1}{2} \sum_{\mu\nu} [x_\mu K_{\mu\nu}^* x_\nu + x_\mu^* K_{\mu\nu} x_\nu^* + 2x_\mu^* (D_{\mu\nu} + L_{\mu\nu}) x_\nu] \\ &= \frac{1}{2} [(x^*)^H K^* x + x^H K x^* + x^H (D + L) x + (x^*)^H (D^* + L^*) x^*] \\ &= \frac{1}{2} z^H \begin{pmatrix} D + L & K \\ K^* & D^* + L^* \end{pmatrix} z, \quad z = \begin{pmatrix} x \\ x^* \end{pmatrix}. \end{aligned} \quad (5.23)$$

The inner matrix is Hermitian, and it can be shown that it can be diagonalized on the form (see Exercise 5.3)

$$\begin{pmatrix} D + L & K \\ K^* & D^* + L^* \end{pmatrix} = \sum_k \lambda_k \begin{pmatrix} u_k \\ u_k^* \end{pmatrix} \begin{pmatrix} u_k \\ u_k^* \end{pmatrix}^H. \quad (5.24)$$

Thus,  $\Delta E_{\text{HF}} \geq 0$  for all  $x$  if and only if all the eigenvalues  $\lambda_k \geq 0$  for this matrix.

### 5.3.1 Exercises

**Exercise 5.1.** Let  $|\Phi\rangle$  be a Slater determinant, and  $\hat{H}$  be a Hamiltonian. Let  $\hat{X} = \sum_{pq} x_{pq} c_p^\dagger c_q$ , with  $x_{pq}$  arbitrary numbers. Show that

$$\langle \Phi | [\hat{H}, \hat{X}] | \Phi \rangle = 0 \quad (5.25)$$

for all  $\hat{X}$  if and only if the occupied orbitals in  $|\Phi\rangle$  solve the HF equations (on non-canonical form). Also show that it is sufficient to conclude this if  $x$  is a unitary matrix.  $\triangle$

**Exercise 5.2.** Do a detailed calculation of

$$A^{(2)} = \langle \Phi | \{\hat{W}\} \{\hat{X}\} \{\hat{X}\} | \Phi \rangle, \quad (5.26)$$

including computing all non-vanishing fully contracted terms. Next, compute

$$(A')^{(2)} = \langle \Phi | \{\hat{X}\} \{\hat{X}\} \{\hat{W}\} | \Phi \rangle. \quad (5.27)$$

Finally, compute

$$B^{(1)} = \langle \Phi | \{\hat{X}\} \{\hat{F}\} \{\hat{X}\} | \Phi \rangle \quad (5.28)$$

and

$$B^{(2)} = \langle \Phi | \{\hat{X}\} \{\hat{W}\} \{\hat{X}\} | \Phi \rangle. \quad (5.29)$$

See the text for the definition of the operators and the correct results.  $\triangle$

**Exercise 5.3.** This is a mathematical exercise of medium difficulty. In the future, it may be rewritten with more details.

Let  $U = U^H$ ,  $V = V^T$  be complex  $n \times n$  matrices, and consider the matrix eigenvalue problem

$$Hz = \begin{pmatrix} U & V \\ V^* & U^* \end{pmatrix} z = \lambda z. \quad (5.30)$$

Show that  $\lambda$  is real, and that the eigenvectors can be chosen on the form  $z = \begin{pmatrix} u \\ u^* \end{pmatrix}$ , where  $u$  is an  $n$ -vector.

Suggested approach: It can be useful to define the *anti-linear* operator  $C$ , that complex-conjugates and flips the two components,

$$Cz = C \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} v^* \\ u^* \end{pmatrix} = Rz^*, \quad R = \begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \quad (5.31)$$

Begin with showing that  $Cz = \alpha z$  if and only if  $\alpha = 1$  and  $z = (u \ u^*)^T$ . Show that  $HC = CH$ , and show that for any eigenvector  $z_k$  of  $H$ ,  $Cz_k$  is also an eigenvector of  $H$  with the same eigenvalue. Let  $z_k$ ,  $k = 1, \dots, n$  be the eigenspace belonging to the eigenvalue  $\lambda$  of  $H$ . Compute the matrix of  $C$  relative to this basis, and show that there is a unitary transformation  $z_k \mapsto \hat{z}_k = \sum_\ell z_\ell U_{\ell k}$  such that the new eigenbasis has  $C = I_n$ . Conclude that you have found an eigenbasis for  $\lambda$  on the form  $z_k = (u_k \ u_k^*)^T$ .  $\triangle$



## Chapter 6

# The Standard Methods of approximation

### 6.1 Introduction

Having dealt with the basic formalism of many-fermion theory, how do we solve the Schrödinger equation approximately? In this section, we discuss the *variational principle*, perhaps *the* most important tool for devising approximate schemes.

We then develop the configuration-interaction method, and then Hartree–Fock theory, and then we combine the two methods.

## 6.2 The variational principle (Not lectured 2017)

Consider the time-independent Schrödinger equation for an  $N$ -fermion system, i.e., given our Hamiltonian  $\hat{H}$ , find a nonzero  $|\Psi\rangle \in L_N^2$  with  $E$  a real number such that

$$\hat{H} |\Psi\rangle = E |\Psi\rangle. \quad (6.1)$$

This is an eigenvalue problem for a Hermitian operator  $\hat{H}$  over a Hilbert space. The mathematical analysis of this problem is complex. However, if the Hilbert space  $L_N^2$  has *finite dimension*  $D$ , then  $\hat{H}$  can be viewed as a *Hermitian matrix*, and we can find a complete set of orthonormal eigenfunctions  $|\Psi_k\rangle$ ,  $k = 0, 1, 2, \dots$  with corresponding eigenvalues  $E_k$ , such that

$$\hat{H} = \sum_{k=0}^D E_k |\Psi_k\rangle \langle \Psi_k|. \quad (6.2)$$

Of course, Hilbert space is usually infinite dimensional, complicating the mathematical analysis of the problem. It may happen that  $\hat{H}$  does not even have a ground state, or not even a single eigenvector. However, it turns out, that in most interesting cases the differences are small enough to warrant the assumption that we are dealing with a finite-dimensional problem, or at least that Eq. (6.2) holds with possibly an infinite dimension.

**Theorem 6.1** (Variational principle). *Consider the expectation value functional defined by*

$$\mathcal{E}(|\Psi\rangle) \equiv \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (6.3)$$

*Let  $|\Psi_*\rangle$  be given. Then  $E_* = \mathcal{E}(|\Psi_*\rangle)$  is a stationary value of  $\mathcal{E}$  with respect to all infinitesimal variations  $|\Psi_*\rangle + \epsilon |\eta\rangle$  (with  $\epsilon$  a small number and  $\langle \eta | \eta \rangle = 1$ ) if and only if*

$$\hat{H} |\Psi_*\rangle = E_* |\Psi_*\rangle. \quad (6.4)$$

*Proof.* Let  $\epsilon$  be a small real number,  $|\Psi\rangle, |\eta\rangle \in L_N^2$  arbitrary vectors,  $|\eta\rangle$  normalized. Let  $f(\epsilon)$  be defined as

$$f(\epsilon) = \mathcal{E}(|\Psi\rangle + \epsilon |\eta\rangle). \quad (6.5)$$

The stationary point condition can be formulated as

$$f'(0) = 0. \quad (6.6)$$

This condition must hold for all  $|\eta\rangle$ . Thus,  $\epsilon |\eta\rangle$  is an arbitrary infinitesimal variation. Mathematically,  $f'(\epsilon)$  is the directional derivative of  $\mathcal{E}$  at  $|\Psi\rangle$  in the direction  $|\eta\rangle$ . Then

$$f(\epsilon) = \frac{\langle \Psi | \hat{H} | \Psi \rangle + \epsilon \langle \eta | \hat{H} | \Psi \rangle + \epsilon \langle \Psi | \hat{H} | \eta \rangle + \epsilon^2 \langle \eta | \hat{H} | \eta \rangle}{\langle \Psi | \Psi \rangle + \epsilon \langle \eta | \Psi \rangle + \epsilon \langle \Psi | \eta \rangle + \epsilon^2 \langle \eta | \eta \rangle}. \quad (6.7)$$

Define  $E = \langle \Psi | \hat{H} | \Psi \rangle$ ,  $N = \langle \Psi | \Psi \rangle$ . Define  $A = \langle \eta | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H} | \eta \rangle$ ,  $a = \langle \eta | \Psi \rangle + \langle \Psi | \eta \rangle$ .

$$\mathcal{E}(|\Psi\rangle + \epsilon |\eta\rangle) = \frac{E + \epsilon A + O(\epsilon^2)}{N + \epsilon a + O(\epsilon^2)}. \quad (6.8)$$

Using  $1/(1+x) = 1 - x + O(x^2)$ , we expand the denominator to first order in  $\epsilon$ :

$$\frac{1}{N + \epsilon a + O(\epsilon^2)} = \frac{1}{N} \left[ 1 - \epsilon \frac{a}{N} + O(\epsilon^2) \right]. \quad (6.9)$$

We expand  $f(\epsilon)$  to first order in  $\epsilon$ :

$$\begin{aligned} Nf(\epsilon) &= \left( E + \epsilon A + O(\epsilon^2) \right) \left[ 1 - \epsilon \frac{a}{N} + O(\epsilon^2) \right] \\ &= E + \epsilon \left[ A - \frac{aE}{N} \right] + O(\epsilon^2). \end{aligned} \quad (6.10)$$

Recall that

$$f(\epsilon) = f(0) + \epsilon f'(0) + O(\epsilon^2). \quad (6.11)$$

We see that  $f'(0) = 0$  if and only if

$$A = \frac{aE}{N}, \quad (6.12)$$

that is,

$$\langle \eta | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H} | \eta \rangle = (\langle \eta | \Psi \rangle + \langle \Psi | \eta \rangle) \mathcal{E}(|\Psi\rangle), \quad (6.13)$$

which must hold for all  $|\eta\rangle$ . In particular, if it holds for  $|\eta\rangle = |u\rangle$  it must also hold for  $|\eta\rangle = i|u\rangle$ . Plugging these in gives

$$\langle u | \hat{H} | \Psi \rangle + \langle \Psi | \hat{H} | u \rangle = (\langle u | \Psi \rangle + \langle \Psi | u \rangle) \mathcal{E}(|\Psi\rangle), \quad (6.14)$$

$$-i \langle u | \hat{H} | \Psi \rangle + i \langle \Psi | \hat{H} | u \rangle = (-i \langle u | \Psi \rangle + i \langle \Psi | u \rangle) \mathcal{E}(|\Psi\rangle), \quad (6.15)$$

Multiplying the second equation by  $i$  and adding the two equations gives

$$\langle u | \hat{H} | \Psi \rangle = \mathcal{E}(|\Psi\rangle) \langle u | \Psi \rangle. \quad (6.16)$$

Since  $|u\rangle$  was arbitrary, we must have

$$\hat{H} | \Psi \rangle = \mathcal{E}(|\Psi\rangle) | \Psi \rangle. \quad (6.17)$$

The proof is complete.  $\square$

The *variational principle* in its simplest form states that the ground-state energy  $E_0$  is the minimum of the *expectation value* of the Hamiltonian:

**Theorem 6.2** (Variational Principle, Rayleigh–Ritz). *If  $\hat{H}$  has a ground state, then the ground-state energy is given by the minimum of the expectation value of  $\hat{H}$ , viz,*

$$E_0 = \min \left\{ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \mid 0 \neq |\Psi\rangle \in L_N^2, \mid \langle \Psi | \hat{H} | \Psi \rangle \mid < +\infty \right\}. \quad (6.18)$$

Theorem 6.2 holds even if Eq. (6.2) does not hold. It is sufficient that  $\hat{H}$  has a lowest eigenvalue. In the infinite dimensional case, we must require that  $\mid \langle \Psi | \hat{H} | \Psi \rangle \mid < +\infty$ , since for most Hamiltonians of interest, there are in fact  $|\Psi\rangle$  that has an infinite expectation value. In finite dimensions, this is of course not true.

We will not prove Theorem 6.2 in its full generality, but we see immediately that it follows from Theorem 6.1:  $E_0$  is a stationary value, and clearly  $\mathcal{E}$  cannot take values *lower* than  $E_0$ . Thus,  $E_0$  must be the minimum.

We now consider *the variational procedure*, a useful method of generating approximate ground-state energies. Suppose we have a *subset of Hilbert space*  $\mathcal{M} \subset L_N^2$ , and compute

$$E_0[\mathcal{M}] \equiv \inf \left\{ \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \mid 0 \neq |\Psi\rangle \in \mathcal{M}, \mid \langle \Psi | \hat{H} | \Psi \rangle \mid < +\infty \right\}. \quad (6.19)$$

Clearly,

$$E_0 \leq E_0[\mathcal{M}], \quad (6.20)$$

since we minimize over a *smaller* set than the full Hilbert space. This *upper bound property* of the variational procedure is very useful, because if we enlarge  $\mathcal{M}$ , we will always get a better estimate for  $E_0$ .

Suppose that our variational procedure yields a minimum value in Eq. (6.19) for the function  $|\tilde{\Psi}\rangle \in \mathcal{M}$ :

$$E_0[\mathcal{M}] = \mathcal{E}(|\tilde{\Psi}\rangle) = \frac{\langle \tilde{\Psi} | \hat{H} | \tilde{\Psi} \rangle}{\langle \tilde{\Psi} | \tilde{\Psi} \rangle}. \quad (6.21)$$

Suppose also that  $|\tilde{\Psi}\rangle$  is fairly close to  $|\Psi_0\rangle$ , i.e.,

$$|\Psi_0\rangle \approx |\tilde{\Psi}\rangle + \epsilon |\eta\rangle \quad (6.22)$$

Then, from the proof of the variational principle, we expect that

$$\tilde{E}_0 - E_0 = f(\epsilon) - f(0) = [f(0) + \epsilon f'(0) + O(\epsilon^2)] - f(0) = O(\epsilon^2), \quad (6.23)$$

i.e., that the error in the eigenvalue is *quadratic* in the error in the eigenfunction! Thus, the error  $E_0[\mathcal{M}] - E_0$  is insensitive to errors in the wavefunction. This explains why the variational procedure is so useful.

Example: The hydrogen atom with Hamiltonian

$$\hat{h} = -\frac{1}{2}\nabla^2 + \frac{1}{r}. \quad (6.24)$$

The exact ground-state wavefunction is well-known,

$$\psi_0(\vec{r}) = C e^{-r}, \quad (6.25)$$

with eigenvalue  $E_0 = -1/2$ . Here,  $C$  is a normalization constant. Let us imagine we did not know  $\psi_0$ , and try a parameterized wavefunction on the form

$$\psi^\alpha(\vec{r}) = (\alpha/\pi)^{3/4} e^{-\alpha r^2/2}. \quad (6.26)$$

Thus,  $\mathcal{M} = \{|\psi^\alpha\rangle \mid \alpha > 0\}$  is the set of approximate wavefunctions, which all satisfy  $\langle \psi^\alpha | \psi^\alpha \rangle = 1$ . We can compute the expectation value,

$$\mathcal{E}(|\psi^\alpha\rangle) = \langle \psi^\alpha | \hat{h} | \psi^\alpha \rangle = \frac{3}{4}\alpha - 2\left(\frac{\alpha}{\pi}\right)^{1/2} \quad (6.27)$$

and minimize with respect to  $\alpha$ ,

$$E_0[\mathcal{M}] = \inf_{\alpha} \mathcal{E}(|\psi^\alpha\rangle) = \mathcal{E}(|\psi^{\frac{16}{9\pi}}\rangle) = -\frac{4}{3\pi} \approx -0.42. \quad (6.28)$$

This is actually a minimum, obtained at  $\alpha = 16/(9\pi)$ . Comparing with the exact result, we see that the energies are rather close for such a simple parameterization. The wavefunctions are not that close, see Fig. 6.1! Note that the exact ground-state is not smooth at  $\vec{r} = 0$ .

Usually, the set  $\mathcal{M}$  contains wavefunction ansätze that are parameterized in some way. In the example, we had a simple Gaussian wavefunction parameterized by the width.

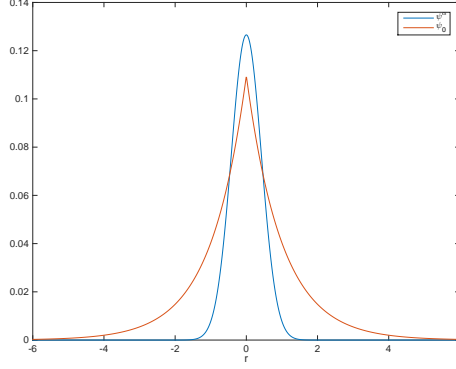


Figure 6.1: Plot of approximate and exact ground-state wavefunction for the Hydrogen example

### 6.2.1 The Cauchy interlace theorem and linear models

Suppose that the set  $\mathcal{M}$  is a linear space, i.e., a subspace  $\mathcal{V}$  of  $L_N^2$  defined by a basis set  $|\Phi_I\rangle$ ,  $I = 1, 2, \dots, D$ . Then the variational procedure is equivalent to computing the smallest eigenvalue of the matrix

$$H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle. \quad (6.29)$$

This is so, because for

$$|\tilde{\Psi}\rangle = \sum_{I=1}^D A_I |\Phi_I\rangle \quad (6.30)$$

the expectation value becomes

$$\mathcal{E}(|\tilde{\Psi}\rangle) = \frac{\mathbf{A}^H \mathbf{H} \mathbf{A}}{\mathbf{A}^H \mathbf{A}}, \quad (6.31)$$

which is simply the expectation value functional for the quantum system with the Hamiltonian  $H$  and wavefunction  $A$ , and we can apply the variational principle to this functional.

It is a fact, that under very mild assumptions on  $\{|\Phi_I\rangle\}$  and  $\hat{H}$ , the eigenvalues of the matrix  $H$  converge to the eigenvalues of  $\hat{H}$ , even in the infinite dimensional case.

For the finite-dimensional case, the Cauchy interlace theorem states that for a linear model as here described, *all* the eigenvalues of  $H$  actually approximate eigenvalues of the full Hamiltonian  $H$  from above. For a general nonlinear model  $\mathcal{M}$ , we cannot say this. In general *only* the ground-state energy is approximated.

The theorem implies that *truncating a single-particle basis* or *truncating a Slater determinant basis* makes sense.

We will not prove the theorem.

**Theorem 6.3** (Cauchy Interlace Theorem). *Let  $\mathcal{V}_1$  and  $\mathcal{V}_2$  be linear spaces, of dimension  $D_1$  and  $D_2$ , respectively. Let  $\mathcal{V}_1 \subset \mathcal{V}_2$  be a subspace.*

*Let  $\{|\Phi_I\rangle\}_{I=1}^{D_2}$  be an orthonormal basis for  $\mathcal{V}_2$ , such that  $\{|\Phi_I\rangle\}_{I=1}^{D_1}$  is a basis for  $\mathcal{V}_1$ .*

*Let  $\hat{H} : \mathcal{V}_2 \rightarrow \mathcal{V}_2$  be a Hermitian operator with matrix  $H_2 \in \mathbb{C}^{D_2 \times D_2}$ ,  $H_{IJ} = \langle \Phi_I | \hat{H} | \Phi_J \rangle$ .*

*Let  $H_1$  be the projection of  $\hat{H}$  onto  $\mathcal{V}_1$ , i.e., the matrix  $H_1$  of this operator is equal to the upper left  $D_1 \times D_1$  block of the  $D_2 \times D_2$  matrix  $H_2$ .*

*Let  $E_k^{(i)}$  be the  $D_i$  eigenvalues of  $H_i$ , arranged such that*

$$E_k^{(i)} \leq E_{k+1}^{(i)} \quad \forall k. \quad (6.32)$$

Then,

$$E_k^{(2)} \leq E_k^{(1)} \leq E_{k+\delta}^{(2)}, \quad \delta = D_2 - D_1. \quad (6.33)$$

### 6.2.2 Exercises

**Exercise 6.1.** Prove Eq. (6.31).

△

## 6.3 The Configuration-interaction method (CI)

### 6.3.1 General description

We now describe an approach to manybody theory called *configuration-interaction theory* (CI). It basically entails truncating both the single-particle basis and the resulting Slater determinant basis according to certain rules.

Let an orthonormal single-particle basis  $\{\phi_p\}$  be given, with associated creation operators  $c_p^\dagger$ , and corresponding Slater determinants  $|\vec{p}\rangle$ . Suppose we expand an  $N$ -fermion wavefunction in the Slater determinant basis, but *truncate* the expansion, including only a finite subset  $\mathcal{S}$  of Slater determinants. The determinants then span a  $D$ -dimensional subspace of  $L_N^2$ ,

$$\mathcal{V} = \text{span}\{|\vec{p}\rangle \mid |\vec{p}\rangle \in \mathcal{S}\} \quad (6.34)$$

Equivalently, any wavefunction in  $V$  can be written

$$|\Psi\rangle = \sum_{\vec{p} \in \mathcal{S}} A_{\vec{p}} |\vec{p}\rangle, \quad A_{\vec{p}} = \langle \vec{p} | \Psi \rangle. \quad (6.35)$$

The set  $\mathcal{S}$  may of course be chosen in many different ways. One typical choice is the set of all possible Slater determinants generated by the first  $L$  single-particle functions  $\phi_0$  through  $\phi_{L-1}$ . This gives a space of dimension  $\binom{L}{N}$ , and is called the *full configuration-interaction space* (FCI space).

Another typical approach is to have a *reference determinant*  $|\Phi\rangle$  and consider particle-hole states on top of that, or excitations in chemistry language.

For example, the one-particle-one-hole space (CI singles, CIS) wavefunction is given by the choice

$$\mathcal{V}_{\text{CIS}} = \text{span}\{|\Phi\rangle, |\Phi_i^a\rangle \mid i = 1, \dots, N, a = N+1, \dots, L\}, \quad (6.36)$$

and any CIS wavefunction can thus be written

$$|\Psi\rangle = A_0 |\Phi\rangle + \sum_{ia} A_i^a |\Phi_{ia}^a\rangle. \quad (6.37)$$

Furthermore, CI singles-and-doubles (CISD) is defined by the space

$$\mathcal{V}_{\text{CISD}} = \text{span}\{|\Phi\rangle, |\Phi_i^a\rangle, |\Phi_{ij}^{ab}\rangle \mid i, j = 1, \dots, N, a, b = N+1, \dots, L\}. \quad (6.38)$$

A wavefunction  $|\Psi\rangle \in \mathcal{V}_{\text{CISD}}$  can be written

$$|\Psi\rangle = A_0 |\Phi\rangle + \sum_{ia} A_i^a |\Phi_i^a\rangle + \sum_{i<j} \sum_{a<b} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle. \quad (6.39)$$

Configuration-interaction singles-doubles-and-triples (CISDT), etc, are defined similarly.

Sometimes, the doubles term is written

$$\sum_{i<j} \sum_{a<b} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle = \frac{1}{4} \sum_{ij} \sum_{ab} A_{ij}^{ab} |\Phi_{ij}^{ab}\rangle. \quad (6.40)$$

The coefficients satisfy  $A_{ij}^{ab} = -A_{ji}^{ab} = -A_{ij}^{ba} = A_{ji}^{ba}$ , and the factor 1/4 comes from the fact that  $|\Phi_{ij}^{ab}\rangle = -|\Phi_{ji}^{ba}\rangle = -|\Phi_{ji}^{ab}\rangle = |\Phi_{ji}^{ba}\rangle$ , i.e., we are deliberately over-counting the basis in this expression to keep notation simple.

Clearly, indexing the Slater determinants using the vector  $\vec{p}$  directly can be cumbersome. Using a different notation, we let  $I \in \mathcal{I}$  be an index that enumerates the basis determinants, and write

$$\mathcal{V} = \text{span}\{|\Phi_I\rangle \mid I \in \mathcal{I}\}. \quad (6.41)$$

Our vector expansion becomes

$$|\Psi\rangle = \sum_I A_I |\Phi_I\rangle, \quad A_I = \langle\Phi_I|\Psi\rangle. \quad (6.42)$$

For example,  $\mathcal{I} = 1, 2, \dots, D$  is a possibility, with some way of choosing an  $I$  for every  $\vec{p}$  we are interested in. Or  $I = (a, i)$ ,  $I = (ab, ij)$ , etc, enumerates the CIS, CISD, etc, hierarchy of spaces.

How do we choose the single-particle functions and the reference state in CI theory? The most common choice in chemistry is to employ a basis of *Hartree–Fock spin-orbitals*. This is the topic of Section 3. A more general picture is as follows: if  $\hat{H} = \hat{H}_0 + \hat{W}$ , it is also possible to consider  $\hat{W}$  a perturbation of  $\hat{H}_0$ , assuming that the eigenstates and eigenvalues of  $\hat{H}_0$  are good approximations to those of the full  $\hat{H}$ . (This is also true for the Hartree–Fock paradigm to be considered later.)

Let therefore  $\{\phi_p\}$  be a complete set of eigenfunctions for the single-particle operator  $\hat{h}$  with eigenvalues  $\epsilon_p$  arranged in increasing order. Then, the Slater determinants  $|\vec{p}\rangle$  are eigenstates of the one-body Hamiltonian  $\hat{H}_0 = \sum_i^N \hat{h}(i)$ . Clearly, the determinant

$$|\Phi\rangle = |123 \cdots N\rangle \quad (6.43)$$

is the ground-state wavefunction of  $\hat{H}_0$ , whose second quantized expression is

$$\hat{H}_0 = \sum_p \epsilon_p c_p^\dagger c_p. \quad (6.44)$$

Note, that if the eigenvalues of  $\hat{h}$  are degenerate, then this wavefunction may or may not be unique.

In this picture, the truncated CI scheme as outlined above is a natural approach, since it is reasonable to assume that singles, doubles, etc, will systematically improve upon the “zero-order” wavefunction  $|\Phi\rangle$ .

In the context of a reference function  $|\Phi\rangle$  defined in terms of a zero-order Hamiltonian, such as  $\hat{H}_0$ , it is common to define the *Fermi level*  $\epsilon$  as the energy of the occupied orbital with the highest energy,  $\epsilon_F$ , assuming that all degenerate levels are included. With this terminology,

$$|\Phi\rangle = \left( \prod_{\epsilon_p \leq \epsilon_F} c_p^\dagger \right) |-\rangle, \quad (6.45)$$

for example. Moreover, we say that a *hole* is “below the Fermi level” and a *particle* is “above the Fermi level”. An excitation excites a fermion from below the Fermi level to above the Fermi level. Thus, the index  $N$  is replaced by the one-body *energy* of that level,  $\epsilon_F$ . See Fig. 6.2

Notice that the truncated CI scheme favors the description of the ground-state wavefunction.

### 6.3.2 Matrix elements of the CI method

Having established the parameterization of the approximate wavefunction, a linear space  $\mathcal{V}$ , we turn to the variational principle, which tells us (together with the Cauchy Interlace Theorem that) that the matrix of the Hamiltonian  $\hat{H}$  with respect to the chosen basis is the central object. Diagonalizing this matrix gives us approximations to the ground-state energy and in total  $D$  eigenvalues of the full system.

Thus, in the CI method, we need to diagonalize the matrix  $H = [H_{IJ}]$  given by

$$H_{IJ} = \langle\Phi_I|\hat{H}|\Phi_J\rangle \quad (6.46)$$



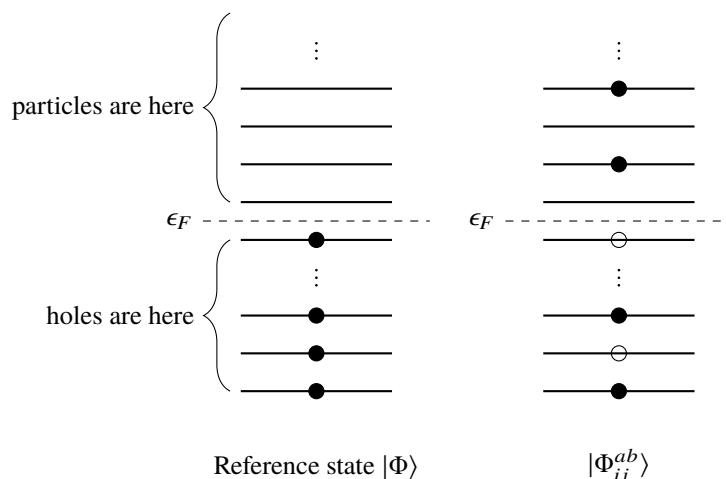


Figure 6.2: Fermi level and quasiparticles. To the left, we have the vacuum state. To the right, we have a doubly excited state, or a two-particle-two-hole-state. Notice how we draw the “Fermi line” between two levels for clarity. In this simple picture, we have assumed that the levels are non-degenerate. If we had spin present, we could fit two particles per level, and so on.

If we look at the CISD case, the matrix then obtains a block form:

$$H = \begin{pmatrix} \langle \Phi | \hat{H} | \Phi \rangle & \langle \Phi | \hat{H} | \Phi_i^a \rangle & \langle \Phi | \hat{H} | \Phi_{ij}^{ab} \rangle \\ \langle \Phi_{i'}^{a'} | \hat{H} | \Phi \rangle & \langle \Phi_{i'}^{a'} | \hat{H} | \Phi_i^a \rangle & \langle \Phi_{i'}^{a'} | \hat{H} | \Phi_{ij}^{ab} \rangle \\ \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi \rangle & \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi_i^a \rangle & \langle \Phi_{i'j'}^{a'b'} | \hat{H} | \Phi_{ij}^{ab} \rangle \end{pmatrix} \quad (6.47)$$

### 6.3.3 Computer implementation of CI methods

In chemistry, *speed* and *reliability* are crucial factors. Computations are performed by non-specialists using highly optimized codes like DALTON, MOLPRO, or GAUSSIAN.

We will not try and compete with such codes, of course, but instead indicate how various methods may be implemented.

### 6.3.4 Naive CI

The simplest approach, which we here call “naive CI”, is to

1. Write down a list of all the Slater determinants in the desired basis,

$$I \mapsto |\Phi_I\rangle.$$

2. Compute all the matrix elements  $H_{IJ}$  and store them in computer memory as a big  $D \times D$  matrix. This can be done using, say, the Slater–Condon rules (see Exercise ??) that are basically formulae for the matrix elements given in terms of the occupied single-particle functions in  $|\Phi_I\rangle$  and  $|\Phi_J\rangle$ .
3. Use a diagonalization algorithm to find, say, the ground-state energy or other eigenvalues of the matrix.

The biggest problem with this approach, is that the dimension  $D$  of the CI space grows pretty fast. The matrix is, in principle, a table with  $D^2$  elements. For FCI,  $D$  grows like  $\binom{L}{N}$ , which very quickly is prohibitive. For CIS, it grows only like  $N(L - N)$ , but CIS is not that fancy. For CISD, the dimension grows like  $N^2(L - N)^2$ . We see that the spaces in any case become huge for moderate particle numbers and numbers  $L$  of single-particle functions.

### 6.3.5 Direct CI

More common than “naive CI” is direct CI. For systems of interest, the matrix size grows so quickly that storing the matrix  $H$  in memory is out of question. Moreover, diagonalization of dense matrices scales as  $D^3$ , quickly becoming too expensive for practical calculations.

Luckily, we have *iterative algorithms* such as the Lanczos algorithm. These rely only on the *matrix-vector product*. Nowhere is the actual value of  $H_{IJ}$  needed, only the action on a vector  $A_I$ , i.e., the algorithm needs to compute

$$\vec{A}' = H\vec{A} \quad (6.48)$$

for some input vector  $\vec{A}$ . I.e., we must have an algorithm to compute

$$|\Psi'\rangle = P\hat{H}|\Psi\rangle \quad (6.49)$$

where  $P = \sum_I |\Phi_I\rangle \langle \Phi_I|$  is the projection operator onto our chosen basis, i.e., we throw away the part of  $\hat{H}|\Psi\rangle$  which is not describable in terms of our basis.

It is useful to represent  $|\Phi_I\rangle$  in terms of its occupation number vector, a bit string  $B = B[I]$ . These are integers, and we need a table of these in computer memory. Since our  $|\Phi_I\rangle$  must be linearly independent, there is a one-to-one correspondence between the  $B[I]$ 's and the  $I$ 's, i.e., we can *invert* the table to obtain  $I = I[B]$ , given  $B$ . We write  $|B\rangle = |\Phi_{I[B]}\rangle$  for brevity, and we stress that now  $B$  is an integer written on binary form.

The central observation is now that, for any string of creation and annihilation operators

$$C_1 C_2 \cdots C_n |B\rangle = \begin{cases} 0 \\ (-1)^s |B'\rangle \end{cases} \quad \text{or} \quad (6.50)$$

The result can be found by *manipulating the bits of  $B$  and keeping track of the resulting sign*. When  $B'$  has been found, the corresponding index  $I'$  can be found by searching the bit pattern table. Thus, let us write:

$$\begin{aligned} |\Psi'\rangle &= P\hat{H}|\Psi\rangle = \sum_{B'} |B'\rangle \langle B'| \hat{H} \sum_B A_B |B\rangle \\ &= \sum_B A_B \sum_{B'} |B'\rangle \langle B'| \left( \sum_{pq} h_q^p c_p^\dagger c_q + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} c_p^\dagger c_q^\dagger c_s c_r \right) |B\rangle \\ &= |\Psi'\rangle = \sum_B A_B \left( \sum_{pq} h_q^p \langle B'| c_p^\dagger c_q |B\rangle + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle B'| c_p^\dagger c_q^\dagger c_s c_r |B\rangle \right) |B'\rangle \end{aligned} \quad (6.51)$$

This gives us the following algorithm for computing the  $\hat{H}_0$  contribution to  $|\Psi'\rangle$  (the  $\hat{W}$  part is similar):

1. Initialize  $A'_{I'} = 0$  for all  $I'$ .
2. Loop over  $I$ :
  - (a) Fetch  $B = B[I]$ .

(b) Loop over  $p, q$ .

- i. Compute  $c_p^\dagger c_q |B\rangle = 0$  or  $(-1)^s |B'\rangle$  by manipulating the bits in  $B$ .
- ii. If the result is nonzero, compute  $I'$  such that  $B[I'] = B'$  by searching the bit pattern table.
- iii. If the pattern is found, update  $A'_{I'} \leftarrow A'_{I'} + A_I h_q^p (-1)^s$ .

Of course, this algorithm is just a sketch. There are many ways to improve it.

How does one search for the index  $I'$  in step 2/b/ii? One way is to ensure that the table of bit patterns (integers) are sorted, and then use *binary search*. This requires on average  $O(D \log D)$  operations, and since we need to do this  $O(D)$  times, this slows down our program drastically. One can also use a *hash map* (e.g., the C++ STL class `std::map<int, int>` can be used). This is no faster.

A *much* faster approach can be taken using *graphical methods*. It is actually possible to find a formula for the inverse map. This formula is  $O(1)$ , dramatically reducing the computer work for direct CI. For more information on this technique, see Helgaker/Jørgensen/Olsen [6], Section 11.8.

### 6.3.6 Recipe for bit pattern representation.

How can we perform the bitwise operations mentioned above?

Each Slater determinant  $|\mu_1, \dots, \mu_N\rangle$  is, via the occupation numbers, maps to the bit pattern  $|n_0 n_1 n_2 \dots n_L\rangle$  where each  $n_\mu \in \{0, 1\}$ . We identify the bit pattern with the integer  $B[\mu_1 \dots \mu_N]$  it encodes. Thus,

$$\vec{\mu} = \{1, 5, 6\} \mapsto \underbrace{010001100 \dots 0_2}_{L \text{ bits}} \mapsto |1 \times 2^1 + 1 \times 2^5 + 1 \times 2^6\rangle = |97_{10}\rangle. \quad (6.52)$$

(But who is thinking in terms of base-10 numbers these days anyway?) All integers between 0 and  $2^{L-1}$  encode all possible Fock space basis functions. A basis for  $N$ -fermion space is composed of all the integers whose bit patterns have precisely  $N$  bits in total.

Annihilation operator:  $c_p |B\rangle$  is either 0 or  $(-1)^k |B'\rangle$  for some  $k$  and  $B'$ . We have the following algorithm:

1. If bit  $p$  is not set, return the zero result.
2. Else, compute  $k$  as the number of bits set *before*  $p$ .
3. Erase bit  $p$  to obtain  $B'$ .
4. Return the sign  $(-1)^k$  and  $B'$ .

Creation operator:  $c_p^\dagger |B\rangle$  is either 0 or  $(-1)^k |B'\rangle$  for some  $k$  and  $B'$ . We have the following algorithm:

1. If bit  $p$  is set, return the zero result.
2. Else, compute  $k$  as the number of bits set *before*  $p$ .
3. Light bit  $p$  to obtain  $B'$ .
4. Return the sign  $(-1)^k$  and  $B'$ .

The product  $c_p^\dagger c_q |B\rangle$  can be computed by repeating the above algorithms, and similarly with *any* string of creation and annihilation operators.

### 6.3.7 Exercises

**Exercise 6.2.** Compute the dimension of  $\mathcal{V}_{\text{CIS}}$ ,  $\mathcal{V}_{\text{CISD}}$ , etc.  $\Delta$

**Exercise 6.3.** We are given  $L = 8$  single-particle functions, numbered  $p = 0, 1, \dots, L - 1$ , and thus an occupation number representation of length 8 bits, e.g.,

$$|p = 2, p = 3\rangle = |0_0 0_1 1_2 1_3 0_4 0_5 0_6 0_7\rangle = |00110000\rangle. \quad (6.53)$$

Write down the result of the following expressions, on occupation number form. Remember the sign factor:

- a)  $c_1^\dagger |01100000\rangle$
- b)  $c_5 |01000101\rangle$
- c)  $c_4^\dagger c_1^\dagger |01101000\rangle$
- d)  $c_1^\dagger |01100000\rangle$
- e)  $c_6 |11111111\rangle$
- f)  $c_6^\dagger |01111001\rangle$
- g)  $c_1^\dagger c_2^\dagger c_3^\dagger |00000000\rangle$
- h)  $c_4^\dagger c_4 |11101000\rangle$

$\Delta$

**Exercise 6.4.** Write a program that generates all possible bit patterns of length  $L$  with  $N$  bits set and writes them to screen.

Check that you have the correct number of patterns,  $\binom{L}{N}$ .

Can you avoid a brute-force loop that checks all the  $2^N$  possibilities? Try a recursive algorithm.  $\Delta$

**Exercise 6.5.** (continues exercise 6.4.) Write a program that correctly creates/annihilates particles from a bit pattern representation  $|B\rangle$  of a Slater determinant, returning the proper sign.  $\Delta$

**Exercise 6.6.** (continues exercises 6.4 and 6.5.) Write a program that, given  $h_q^p$  and  $w_{rs}^{pq}$  (antisymmetrized or otherwise) as input arrays, computes  $\hat{H} |B\rangle$  using direct CI.  $\Delta$

## 6.4 Perturbation theory for the ground-state (PT)

### 6.4.1 Non-degenerate Rayleigh–Schrödinger perturbation theory (RSPT)

Perturbation theory is a powerful method for systematic improvement of a model wavefunction. We can for the moment “forget” everything we know about second quantization, Slater determinants, quasiparticles, etc: PT is a generic theory applicable to all matrix problems.

Supporting material: Szabo and Ostlund; Bartlett and Shavitt; Helgaker, Jørgensen and Olsen.

Suppose we have a Hamiltonian  $\hat{H}$  for which we seek eigenfunctions and eigenvalues,

$$\hat{H} |\Psi_k\rangle = E_k |\Psi_k\rangle. \quad (6.54)$$

The idea is to partition the Hamiltonian into a part that we can “solve” and a perturbation  $\hat{V}$ ,

$$\hat{H} = \hat{H}_0 + \hat{V}. \quad (6.55)$$

The operator  $\hat{H}_0$  is “solved”, in the sense that we assume knowledge of all its eigenfunctions and eigenvalues,

$$\hat{H}_0 |\Phi_k\rangle = \epsilon_k |\Phi_k\rangle. \quad (6.56)$$

The set  $\{|\Phi_k\rangle\}$  is assumed to be an orthonormal basis for Hilbert space (this is true for all finite-dimensional cases, and for many infinite-dimensional ones). We should, in principle, be able to express the exact eigenvectors and (and therefore the eigenvalues) in terms of this basis and  $\hat{V}$ .

In perturbation theory, we seek such an expression in terms of *power series in the perturbation*  $\hat{V}$ . We introduce an *order parameter*  $\lambda$  and write

$$\hat{H}_\lambda = \hat{H}_0 + \lambda \hat{V}, \quad (6.57)$$

i.e.,  $\hat{H} = \hat{H}_1$  is the full Hamiltonian. It is not unreasonable to assume that the eigenvalues and eigenvectors of  $\hat{H}_\lambda$  become smooth functions of  $\lambda$ , at least for  $\lambda$  sufficiently small and/or sufficiently weak perturbations  $\hat{V}$ .

The Schrödinger equation for  $\hat{H}(\lambda)$  reads

$$\hat{H}_\lambda |\Psi_k(\lambda)\rangle = E_k(\lambda) |\Psi_k(\lambda)\rangle. \quad (6.58)$$

We now assume that we can expand the eigenvectors and eigenvalues in *power series* around  $\lambda = 0$ .

$$|\Psi_k(\lambda)\rangle = \sum_{n=0}^{\infty} |\Psi_k^{(n)}\rangle \lambda^n \quad (6.59a)$$

$$E_k(\lambda) = \sum_{n=0}^{\infty} E_k^{(n)} \lambda^n. \quad (6.59b)$$

The unperturbed problem is obtained at  $\lambda = 0$ :  $|\Psi_k(0)\rangle = |\Phi_k\rangle = |\Psi_k^{(0)}\rangle$  and  $E_k(0) = \epsilon_k$ , and the full problem at  $\lambda = 1$ :  $|\Psi_k(1)\rangle = |\Psi_k\rangle$ , and  $E_k(1) = E_k$ .

The assumption that  $|\Psi_k(\lambda)\rangle$  is differentiable at  $\lambda = 0$  is assured by requiring  $\epsilon_k$  to be *non-degenerate*.

We now derive formulas for the perturbation corrections  $E_k^{(n)}$  and  $|\Psi_k^{(n)}\rangle$ . This is done by plugging Eqs. (6.59) into the Schrödinger equation. This gives

$$(\hat{H}_0 + \lambda \hat{V}) \sum_{n=0}^{\infty} |\Psi_k^{(n)}\rangle \lambda^n = \left( \sum_{n=0}^{\infty} E_k^{(n)} \lambda^n \right) \sum_{m=0}^{\infty} |\Psi_k^{(m)}\rangle \lambda^m. \quad (6.60)$$

For this equation to hold *for all*  $\lambda$ , it must hold order-by-order. The  $\lambda^0$ -part of the equation is simply Eq. (6.56). The  $n$ 'th order equation is

$$\hat{H}_0 |\Psi_k^{(n)}\rangle + \hat{V} |\Psi_k^{(n-1)}\rangle = \sum_{j=0}^n E_k^{(j)} |\Psi_k^{(n-j)}\rangle, \quad n > 0. \quad (6.61)$$

The solution  $|\Psi_k(\lambda)\rangle$  to the Schrödinger equation is not unique. By scaling it we obtain a new solution. Thus, in order to write  $|\Psi_k(\lambda)\rangle$  as a smooth function of  $\lambda$ , we need to select one particular normalization for each  $\lambda$ . We obtain particularly simple expressions using *intermediate normalization*:

$$\langle \Phi_k | \Psi_k(\lambda) \rangle = 1. \quad (6.62)$$

Inserting the power series for  $|\Psi_k(\lambda)\rangle$  we obtain

$$1 = \langle \Phi_k | \Psi_k(\lambda) \rangle = 1 + \lambda \langle \Phi_k | \Psi_k^{(1)} \rangle + \lambda^2 \langle \Phi_k | \Psi_k^{(2)} \rangle + \dots, \quad (6.63)$$

Since this expression is to hold for all  $\lambda$ , it must hold order-by-order, which gives

$$\langle \Phi_k | \Psi_k^{(n)} \rangle = 0, \quad \forall n \geq 1, \quad (6.64)$$

i.e., *all* the higher-order corrections are orthogonal to the unperturbed vector  $|\Phi_k\rangle$ .

We now use Eq. (6.64) and project Eq. (6.61) onto  $|\Phi_k\rangle$  to obtain

$$\langle \Phi_k | \hat{V} | \Psi_k^{(n-1)} \rangle = E_k^{(n)}, \quad (6.65)$$

which is an expression for the  $n$ -th order energy perturbation in terms of the  $n - 1$ -th order correction in the wavefunction. In particular,

$$E_k^{(1)} = \langle \Phi_k | \hat{V} | \Phi_k \rangle. \quad (6.66)$$

If we can find an expression for  $|\Psi_k^{(n)}\rangle$  in terms of  $|\Psi_k^{(j)}\rangle$ ,  $j < n$ , then we have a recursive procedure for determining all the perturbation corrections.

To this end, rearrange Eq. (6.61) as

$$(\epsilon_k - \hat{H}_0) |\Psi_k^{(n)}\rangle = \hat{V} |\Psi_k^{(n-1)}\rangle - \sum_{j=0}^{n-1} E_k^{(n-j)} |\Psi_k^{(j)}\rangle. \quad (6.67)$$

On the right-hand side we only have wavefunction corrections of order less than  $n$ . We also know that the  $E_k^{(n)}$ , which occurs on the right-hand side, is a function of  $|\Psi_k^{(n-1)}\rangle$ , so if we can somehow invert  $\epsilon_k - \hat{H}_0$  then we have an expression for  $|\Psi_k^{(n)}\rangle$  in terms of lower-order corrections only.

Let  $\hat{P} = |\Phi_k\rangle \langle \Phi_k|$ , the projection operator onto the unperturbed eigenvector. Let  $\hat{Q} = 1 - \hat{P}$ , which is then the projector onto the subspace spanned by all the other  $|\Phi_j\rangle$ ,  $j \neq k$ :

$$\hat{Q} = \sum_{j \neq k} |\Phi_j\rangle \langle \Phi_j|. \quad (6.68)$$

Intermediate normalization can now be written

$$|\Psi_k^{(n)}\rangle = \hat{Q} |\Psi_k^{(n)}\rangle, \quad n \geq 1. \quad (6.69)$$

Moreover,

$$[\hat{H}_0, \hat{Q}] = 0, \quad (6.70)$$

since the  $|\Phi_j\rangle$  are eigenfunctions of  $\hat{H}_0$ . Acting on Eq. (6.67) with  $\hat{Q}$  we then obtain

$$(\epsilon_k - \hat{H}_0)\hat{Q}|\Psi_k^{(n)}\rangle = \hat{Q}\hat{V}|\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)}\hat{Q}|\Psi_k^{(j)}\rangle, \quad (6.71)$$

where we remark that the  $j = 0$ -term from the sum on the right-hand side is eliminated. We have

$$\epsilon_k - \hat{H}_0 = \sum_{j \neq k} (\epsilon_k - \epsilon_j) |\Phi_j\rangle \langle \Phi_j| = \hat{Q}(\epsilon_k - \hat{H}_0)\hat{Q}, \quad (6.72)$$

i.e., the operator acts only within the space orthogonal to  $|\Phi_k\rangle$ . (Here, we use the non-degeneracy assumption.) Define the operator

$$\hat{R} = \hat{Q}\hat{R}\hat{Q} = \sum_{j, j \neq k} \frac{1}{\epsilon_k - \epsilon_j} |\Phi_j\rangle \langle \Phi_j|. \quad (6.73)$$

It is important to note that we must *here assume that the unperturbed eigenvalue  $\epsilon_k$  is non-degenerate*. Otherwise there are infinite terms in the sum. Now, for every  $|u\rangle = \hat{Q}|u\rangle$  (such as  $|\Psi_k^{(n)}\rangle$ ) we have

$$\hat{R}(\epsilon_k - \hat{H}_0)|u\rangle = |u\rangle. \quad (6.74)$$

The operator  $\hat{R}$  is called a *pseudoinverse*, and since  $\hat{R} = \hat{Q}\hat{R}\hat{Q}$  it is common to write

$$\hat{R} = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0}, \quad (6.75)$$

even though the fraction notation for matrices and operators is something to be careful with.  $\hat{R}$  is also called the *resolvent* of  $\hat{H}_0$ . Acting with  $\hat{R}$  on Eq. (6.67), we obtain

$$|\Psi_k^{(n)}\rangle = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0} \left[ \hat{V}|\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)} |\Psi_k^{(j)}\rangle \right]. \quad (6.76)$$

We summarize as a theorem:

**Theorem 6.4** (Non-degenerate Rayleigh–Schrödinger Perturbation Theory). *Let  $\hat{H} = \hat{H}_0 + \lambda\hat{V}$  be given, and assume that*

$$\hat{H}_0 |\Phi_k\rangle = \epsilon_k |\Phi_k\rangle \quad (6.77)$$

*where the eigenvectors for a complete basis. Let*

$$(\hat{H}_0 + \lambda\hat{V})|\Psi_k(\lambda)\rangle = E_k(\lambda)|\Psi_k(\lambda)\rangle, \quad \langle \Phi_k | \Psi_k(\lambda) \rangle = 1, \quad (6.78)$$

*for a given  $k$ , and assume that  $\epsilon_k$  is a non-degenerate eigenvalue for  $\hat{H}_0$ . Assume furthermore, that  $E_k(\lambda)$  and  $|\Psi_k(\lambda)\rangle$  are analytic in a neighborhood of  $\lambda = 0$ ,*

$$E_k(\lambda) = \sum_{n=0}^{\infty} E_k^{(n)} \lambda^n \quad (6.79)$$

$$|\Psi_k(\lambda)\rangle = \sum_{n=0}^{\infty} |\Psi_k^{(n)}\rangle \lambda^n. \quad (6.80)$$

Then the  $n$ -th order corrections are given recursively in terms of the  $j < n$ -th order corrections via the formulae

$$E_k^{(n)} = \langle \Phi_k | \hat{V} | \Psi_k^{(n-1)} \rangle \quad (6.81)$$

$$|\Psi_k^{(n)}\rangle = \frac{\hat{Q}}{\epsilon_k - \hat{H}_0} \left[ \hat{V} |\Psi_k^{(n-1)}\rangle - \sum_{j=1}^{n-1} E_k^{(n-j)} |\Psi_k^{(j)}\rangle \right]. \quad (6.82)$$

#### 6.4.2 Low-order RSPT

Theorem 6.4 gives a recursive procedure for the  $n$ -th order corrections of the energies and wavefunctions. We now consider the explicit expressions up to  $n = 3$ .

For notational simplicity, we omit the subscript  $k$  in the following, and write  $\epsilon \equiv \epsilon_k$  for the unperturbed energy,  $|\Phi\rangle \equiv |\Phi_k\rangle$  for the unperturbed wavefunction, etc. We use  $\hat{R}$  for the resolvent (which also depends on  $k$ ).

The first-order correction to the energy is simple,

$$E^{(1)} = \langle \Phi | \hat{V} | \Phi \rangle. \quad (6.83)$$

For  $E^{(2)}$  we need the first-order wavefunction correction,

$$|\Psi^{(1)}\rangle = \hat{R} \hat{V} |\Phi\rangle, \quad (6.84)$$

which then gives

$$E^{(2)} = \langle \Phi | \hat{V} \hat{R} \hat{V} | \Phi \rangle = \sum_{j, j \neq k} \frac{|\langle \Phi_k | \hat{V} | \Phi_j \rangle|^2}{\epsilon_k - \epsilon_j}, \quad (6.85)$$

which is a familiar expression for second-order perturbation theory. For  $E^{(3)}$  we need the second-order wavefunction correction,

$$|\Psi^{(2)}\rangle = \hat{R} [\hat{V} - \langle \Phi | \hat{V} | \Phi \rangle] \hat{R} \hat{V} |\Phi\rangle. \quad (6.86)$$

This gives

$$\begin{aligned} E^{(3)} &= \langle \Phi | \hat{V} \hat{R} [\hat{V} - \langle \Phi | \hat{V} | \Phi \rangle] \hat{R} \hat{V} | \Phi \rangle \\ &= \langle \Phi | \hat{V} \hat{R} \hat{V} \hat{R} \hat{V} | \Phi \rangle - \langle \Phi | \hat{V} | \Phi \rangle \langle \Phi | \hat{V} \hat{R}^2 \hat{V} | \Phi \rangle \end{aligned} \quad (6.87)$$

Comparing  $E^{(1)}$ , and  $E^{(2)}$ , we notice a pattern, but that for  $E^{(3)}$ , we see that the pattern becomes more complicated: there is a leading term on the form

$$E_{\text{leading}}^{(k)} = \underbrace{\langle \Phi | \hat{V} \hat{R} \hat{V} \hat{R} \cdots \hat{V} | \Phi \rangle}_{n \text{ factors } \hat{V}, n-1 \text{ factors } \hat{R}}, \quad (6.88)$$

but then there are terms arising from

$$E^{(j)} \langle \Psi^{(n-1)} | \Psi^{(n-j)} \rangle = E^{(j)} \langle \Phi | \hat{V} \text{ monomial}(\hat{R}, \hat{V}) \hat{V} | \Phi \rangle, \quad (6.89)$$

where the monomial is a product of in total  $n - j - 2$  operators  $\hat{V}$ , and several  $\hat{R}$ s, in some order.

We will not consider the perturbation series further here (see Exercise 6.8). However, when we discuss diagrams, the second term of  $E^{(3)}$  is referred to as an unlinked term. Higher order corrections obtain more and more such terms. They can be systematically generated from  $E_{\text{leading}}^{(n)}$  by a procedure called the “bracketing procedure”, see Paldus and Čížek[7] in the supporting material for more details.



The  $n$ th order energy  $E^{(n)}$  can be written as the leading term plus terms generated by inserting  $\langle \rangle$  around one or more  $\hat{V}$ s, except for the outer ones, in any possible way, in any number. They may also be nested. The bra(c)ket represents an expectation value with  $|\Psi\rangle$ . The sign of each term is  $(-1)^j$ , where  $j$  is the number of brackets in the term. For example, for  $n = 3$  there is one possibility:

$$\langle \Phi | \hat{V} \hat{R} \langle \hat{V} \rangle \hat{R} \hat{V} | \Phi \rangle = - \langle \Phi | \hat{V} | \Phi \rangle \langle \Phi | \hat{V} \hat{R}^2 \hat{V} | \Phi \rangle, \quad (6.90)$$

which reproduces  $E^{(3)}$ . For  $n = 4$  we have the possibilities

$$\langle \Phi | \hat{V} \hat{R} \langle \hat{V} \rangle \hat{R} \hat{V} \hat{R} \hat{V} | \Phi \rangle = - \langle \Phi | \hat{V} | \Phi \rangle \langle \Phi | \hat{V} \hat{R}^2 \hat{V} \hat{R} \hat{V} | \Phi \rangle \quad (6.91a)$$

$$\langle \Phi | \hat{V} \hat{R} \hat{V} \hat{R} \langle \hat{V} \rangle \hat{R} \hat{V} | \Phi \rangle = - \langle \Phi | \hat{V} | \Phi \rangle \langle \Phi | \hat{V} \hat{R} \hat{V} \hat{R}^2 \hat{V} | \Phi \rangle \quad (6.91b)$$

$$\langle \Phi | \hat{V} \hat{R} \langle \hat{V} \rangle \hat{R} \langle \hat{V} \rangle \hat{R} \hat{V} | \Phi \rangle = \langle \Phi | \hat{V} | \Phi \rangle^2 \langle \Phi | \hat{V} \hat{R}^3 \hat{V} | \Phi \rangle \quad (6.91c)$$

$$\langle \Phi | \hat{V} \hat{R} \langle \hat{V} \hat{R} \hat{V} \rangle \hat{R} \hat{V} | \Phi \rangle = - \langle \Phi | \hat{V} \hat{R} \hat{V} | \Phi \rangle \langle \Phi | \hat{V} \hat{R}^2 \hat{V} | \Phi \rangle. \quad (6.91d)$$

In higher order energies, one even gets brackets within brackets, and quite a lot of terms.

### 6.4.3 A two-state example

It is instructive to consider a two-state example, since we can diagonalize it exactly and obtain closed-form expressions. The behavior of the perturbation series can then be considered.

Let

$$H_0 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & \epsilon \\ \epsilon & 0 \end{pmatrix}, \quad (6.92)$$

so that

$$H(\lambda) = \begin{pmatrix} -1 & \lambda\epsilon \\ \lambda\epsilon & 1 \end{pmatrix}. \quad (6.93)$$

The exact eigenvalues of  $H(\lambda)$  are given by the roots of the polynomial

$$\det(H(\lambda) - eI) = (-1 - e)(1 - e) - (\lambda\epsilon)^2 = -(1 + e)(1 - e) - (\lambda\epsilon)^2. \quad (6.94)$$

Solving, we find the two roots

$$e_{\pm} = \pm[1 + (\lambda\epsilon)^2]^{1/2}. \quad (6.95)$$

As functions of  $\lambda$ , see Fig. 6.3. Note well, that our calculations are also true for *complex*  $\lambda$ , only that  $e_{\pm}$  are no longer real, but complex roots in general.

In RSPT, we would seek the Taylor series of, say,  $e_-(\lambda)$  around  $\lambda = 0$ . Since we have a closed-form expression we can compute this series. The first three terms are

$$e_-(\lambda) = e_-(0) + \lambda e'_-(0) + \frac{1}{2} \lambda^2 e''_-(0) + O(\lambda^3). \quad (6.96)$$

Explicit evaluation of the derivatives:

$$e'_-(\lambda) = -[1 + (\lambda\epsilon)^2]^{-1/2} \lambda\epsilon^2 \quad (6.97)$$

$$e''_-(\lambda) = [1 + (\lambda\epsilon)^2]^{-3/2} (\lambda\epsilon^2)^2 - [1 + (\lambda\epsilon)^2]^{-1/2} \epsilon^2 \quad (6.98)$$

We obtain the Taylor series (with a few extra terms obtained by computer algebra)

$$e_-(\lambda) = -1 - \frac{1}{2} (\lambda\epsilon)^2 + \frac{1}{8} (\lambda\epsilon)^4 - \frac{1}{26} (\lambda\epsilon)^6 + \frac{5}{128} (\lambda\epsilon)^8 + O(\lambda^{10}). \quad (6.99)$$

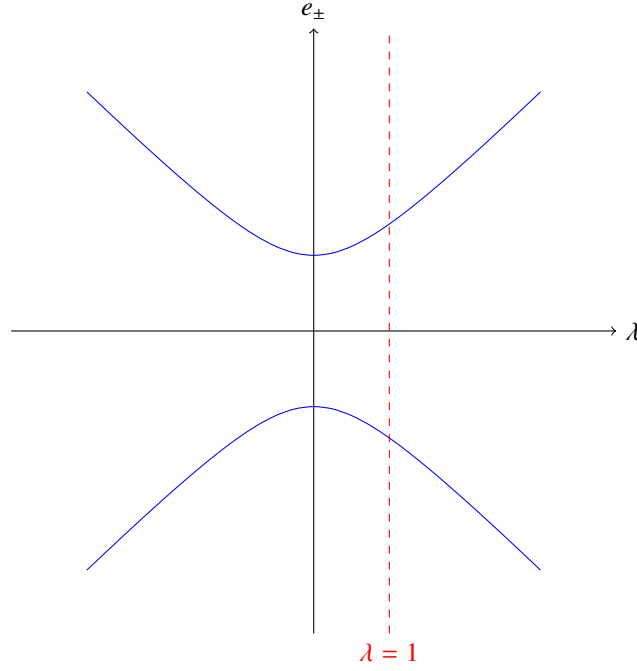


Figure 6.3: The eigenvalues of a two-state problem, as function of the perturbation parameter  $\lambda$ . Here,  $\epsilon = 1$ .

A natural question arises: does the series *converge*? Does it converge for our desired parameter value  $\lambda = 1$ ? Well, our function has a *branch-point singularity* since it is a square-root function. The branch-point singularity arises when the two roots coincide in the complex plane, at  $\lambda = \pm i/\epsilon$ . At these points the eigenvalue functions are no longer analytic. The Taylor series only converges in a disc around  $\lambda = 0$  that does not contain the singularity. Thus, the Taylor series will only converge within the circle  $|\lambda| < 1/|\epsilon|$ , i.e., it will converge for  $\lambda = 1$  only if  $|\epsilon| < 1$ .

Thus, we see directly that the strength of the perturbation may affect the convergence properties of the perturbation series.

The points  $\lambda = \pm i/\epsilon$  are called *avoided crossings* since, if the parameter  $\lambda$  is real, it “narrowly misses” the branch-point and hence an exact crossing. Often, in eigenvalue plots, one can see the function behavior  $\pm[a + (\epsilon\lambda)^2]^{1/2}$ , indicating an avoided crossing and hence a singularity located approximately at this  $\lambda$ -value.

For an  $n$ -state problem, each of the  $n$  eigenvalues may collide with  $n-1$  eigenvalues (again for complex  $\lambda$  in general), giving quite a lot of possible branch points, and thus many singularities. Determining whether the RSPT series converges is thus a virtually impossible task for many-body calculations. Still, a few terms may still give a good approximation.

We now consider the perturbation series of the two-state problem explicitly. We write  $|0\rangle$  for the unperturbed ground state, and  $|1\rangle$  for the unperturbed excited state. We obtain

$$\hat{H}_0 = |0\rangle\langle 0| - |1\rangle\langle 1|, \quad \hat{V} = \epsilon(|0\rangle\langle 1| + |1\rangle\langle 0|). \quad (6.100)$$

We also have

$$\hat{R} = -\frac{1}{2} |2\rangle\langle 2|. \quad (6.101)$$

Perturbation terms for the energy:

$$E^{(1)} = \langle 0 | \hat{V} | 0 \rangle = 0. \quad (6.102)$$

$$E^{(2)} = \langle 0 | \hat{V} \hat{R} \hat{V} | 0 \rangle = -\frac{1}{2} \langle 0 | \hat{V} | 1 \rangle \langle 1 | \hat{V} | 0 \rangle = -\frac{1}{2} \epsilon^2. \quad (6.103)$$

For the higher order terms, we note that  $\hat{R} \hat{V} \hat{R} = 0$ . The third-order energy:

$$E^{(3)} = \langle 0 | \hat{V} \hat{R} \hat{V} \hat{R} \hat{V} | 0 \rangle - \langle 0 | \hat{V} | 0 \rangle \langle 0 | \hat{V} \hat{R}^2 \hat{V} | 0 \rangle = 0. \quad (6.104)$$

#### 6.4.4 Manybody Perturbation Theory (MBPT)

We now use nondegenerate RSPT, and apply it to a manybody Hamiltonian. Thus, manybody Rayleigh-Schrödinger perturbation theory for nondegenerate states.

In most cases, one has a partitioning  $\hat{H} = \hat{K} + \hat{L}$ , where  $\hat{K}$  is a onebody operator (say,  $\hat{H}_0$  or  $\hat{F}$ , the Fock operator), and where  $\hat{L} = \hat{H} - \hat{K}$  may be a two plus onebody operator. Thus,

$$\hat{H} = \hat{K} + \hat{L}, \quad (6.105)$$

$$\hat{K} = \sum_{i=1}^N \hat{k}(i), \quad (6.106)$$

$$\hat{L} = \sum_{i=1}^N \hat{\ell}^{(1)}(i) + \sum_{i<j}^N \hat{\ell}^{(2)}(i, j). \quad (6.107)$$

We take  $\hat{K}$  to be the zero-order Hamiltonian: we assume that  $\hat{k}$  has been diagonalized, giving a complete orthonormal set of single-particle functions,

$$\hat{k} \phi_p(x) = \kappa_p \phi_p(x). \quad (6.108)$$

We introduce a set of creation operators  $c_p^\dagger$  for these functions, and obtain

$$\hat{K} = \sum_p \kappa_p c_p^\dagger c_p \quad (6.109)$$

$$\hat{L} = \sum_{pq} \langle \phi_p | \hat{\ell}^{(1)} | \phi_q \rangle c_p^\dagger c_q + \frac{1}{4} \sum_{pqrs} \langle \phi_p \phi_q | \hat{\ell}^{(2)} | \phi_r \phi_s \rangle_{\text{AS}} c_p^\dagger c_q^\dagger c_s c_r. \quad (6.110)$$

We are considering RSPT for the ground-state wavefunction  $|\Psi\rangle$ . The corresponding unperturbed ground-state is  $|\Phi\rangle = c_1^\dagger \cdots c_N^\dagger |-\rangle$ , the ground-state of  $\hat{K}$ . The unperturbed energy is

$$\epsilon = \langle \Phi | \hat{K} | \Phi \rangle = \sum_{i=1}^N \kappa_i. \quad (6.111)$$

Let us introduce quasiparticle operators, and write  $|\Phi_X\rangle$  for an arbitrary excited Slater determinant. Thus  $|\Phi_X\rangle$  is a Slater determinant with 1 particle-hole pair, 2 particle-hole pairs, etc. We write  $\#X$  for the number of particle-hole pairs in  $|\Phi_X\rangle$ . Thus, the whole complete Slater determinant basis can be constructed:

$$|\Phi_X\rangle = b_{a_1}^\dagger b_{i_1}^\dagger \cdots b_{a_{\#X}}^\dagger b_{i_{\#X}}^\dagger |\Phi\rangle. \quad (6.112)$$

For the projector  $\hat{Q}$ , we get

$$\hat{Q} = \sum_X |\Phi_X\rangle \langle \Phi_X|. \quad (6.113)$$

Note that the reference  $|\Phi\rangle$  is excluded from this sum. The unperturbed energies are

$$\epsilon_X = \langle \Phi_X | \hat{K} | \Phi_X \rangle = \sum_{i=1}^N \kappa_i + \sum_{j=1}^{\#X} (\kappa_{a_j} - \kappa_{i_j}), \quad (6.114)$$

and it is a useful exercise to verify this, see Exercise 6.11. We get for the resolvent

$$\hat{R} = \sum_X \frac{|\Phi_X\rangle \langle \Phi_X|}{\Delta\epsilon_X}, \quad (6.115)$$

where we have defined

$$\Delta\epsilon_X \equiv \epsilon - \epsilon_X = \sum_{j=1}^{\#X} (\kappa_{i_j} - \kappa_{a_j}). \quad (6.116)$$

We here remark that if  $\hat{k}$  has degenerate eigenvalues, we will end up with a zero denominator for some  $X$ . Hence, the assumption of non-degeneracy. In fact, it is not enough to require  $\hat{k}$  to have nondegenerate eigenvalues – we must require that  $\Delta\epsilon_X \neq 0$  for all  $X$ . This is a stronger requirement.

Let us consider RSPT up to second order. We need the matrix elements

$$\langle \Phi_i^a | \hat{L}^{(1)} | \Phi \rangle = \langle \phi_a | \hat{\ell}^{(1)} | \phi_i \rangle \quad (6.117)$$

$$\langle \Phi_i^a | \hat{L}^{(2)} | \Phi \rangle = \sum_j \langle \phi_j \phi_a | \hat{\ell}^{(2)} | \phi_j \phi_i \rangle_{\text{AS}} \quad (6.118)$$

$$\langle \Phi_{ij}^{ab} | \hat{L}^{(2)} | \Phi \rangle = \langle \phi_a \phi_b | \hat{\ell}^{(2)} | \phi_i \phi_j \rangle_{\text{AS}}. \quad (6.119)$$

$$\langle \Phi_i^a | \hat{L}^{(1)} | \Phi \rangle = \langle \phi_a | \hat{\ell}^{(1)} | \phi_i \rangle \quad (6.120)$$

$$\langle \Phi_i^a | \hat{L}^{(2)} | \Phi \rangle = \sum_j \langle \phi_j \phi_a | \hat{\ell}^{(2)} | \phi_j \phi_i \rangle_{\text{AS}} \quad (6.121)$$

$$\langle \Phi_{ij}^{ab} | \hat{L}^{(2)} | \Phi \rangle = \langle \phi_a \phi_b | \hat{\ell}^{(2)} | \phi_i \phi_j \rangle_{\text{AS}}. \quad (6.122)$$

All other matrix elements involving  $\hat{L}$  and  $|\Phi\rangle$  vanish. See Exercise 6.12 for the calculation of the matrix elements.

$$E^{(1)} = \langle \Phi | \hat{L} | \Phi \rangle = \sum_i \langle \phi_i | \hat{\ell}^{(1)} | \phi_i \rangle + \frac{1}{2} \sum_{ij} \langle \phi_i \phi_j | \hat{\ell}^{(2)} | \phi_i \phi_j \rangle_{\text{AS}}. \quad (6.123)$$

$$E^{(2)} = \langle \Phi | \hat{L} \hat{R} \hat{L} | \Phi \rangle = \sum_X \frac{|\langle \Phi_X | \hat{L} | \Phi \rangle|^2}{\Delta\epsilon_X}. \quad (6.124)$$

Since  $\hat{L}$  is at most a two-body operator, this series truncates at  $\#X = 2$ ,

$$\begin{aligned} E^{(2)} &= \sum_{ia} \frac{|\langle \Phi_i^a | \hat{L} | \Phi \rangle|^2}{\kappa_i - \kappa_a} + \frac{1}{4} \sum_{ijab} \frac{|\langle \Phi_{ij}^{ab} | \hat{L}^{(2)} | \Phi \rangle|^2}{\kappa_i + \kappa_j - \kappa_a - \kappa_b} \\ &= \sum_{ia} \frac{|\langle \phi_a | \hat{\ell}^{(1)} | \phi_i \rangle + \sum_j \langle \phi_j \phi_a | \hat{\ell}^{(2)} | \phi_j \phi_i \rangle_{\text{AS}}|^2}{\kappa_i - \kappa_a} + \frac{1}{4} \sum_{ijab} \frac{|\langle \phi_a \phi_b | \hat{\ell}^{(2)} | \phi_i \phi_j \rangle|^2}{\kappa_i + \kappa_j - \kappa_a - \kappa_b} \end{aligned} \quad (6.125)$$

The prefactor comes from over-counting the double excitations. We note that only the two-body part of  $\hat{L}$  contributes in the second sum.

Higher-order corrections quickly become more complicated. We shall see, that using the *Hamiltonian on normal-order form* will simplify matters a lot. Moreover, in Møller–Plesset perturbation theory, where we use  $\hat{K} = \hat{F}$ , the Fock operator, we get even more simplifications due to Brillouin’s Theorem.

### 6.4.5 Møller–Plesset Perturbation Theory

When we partition the Hamiltonian according to HF theory, we obtain *Møller–Plesset PT*. The  $n$ -th order theory is called  $MP_n$ .  $MP2$ – $4$  are standard in most quantum chemical implementations, but not higher. The cost of  $MP4$  is quite high, comparable to the coupled-cluster with singles, doubles, and perturbative triples (CCSD(T), see later in the course). The latter is more accurate in most cases. The perturbation theory nature of MP means that we have in general no way of knowing if the series will converge, and therefore whether or not the  $MP_n$  approximation is good.

For a different account of the present material, see Szabo/Ostlund section 6.5 (p. 350).

We partition the *normal-ordered* Hamiltonian as

$$\hat{H}_N = \hat{H} - \hat{E}_{\text{HF}} = \{\hat{F}\} + \{\hat{W}\}. \quad (6.126)$$

Working with a normal-ordered partitioning simplifies things a little. We now assume that we work in a basis where  $\hat{F}$  is diagonal,

$$\hat{F} |\Phi\rangle = e_0 |\Phi\rangle, \quad \hat{F} |\Phi_X\rangle = e_X |\Phi_X\rangle. \quad (6.127)$$

Here  $X = \binom{A_1 \dots A_n}{I_1 \dots I_n}$  is an excitation index. Now, we have

$$\{\hat{F}\} = \hat{F} - \langle \Phi | \hat{F} | \Phi \rangle, \quad (6.128)$$

and obtain  $\{\hat{F}\} |\Phi\rangle = 0$ , and that

$$\{\hat{F}\} |\Phi_X\rangle = (e_X - e_0) |\Phi_X\rangle = \left( \sum_j \epsilon_{a_j} - \epsilon_{i_j} \right) |\Phi_X\rangle. \quad (6.129)$$

Thus, we seek the perturbation expansion

$$E_0 = E_{\text{HF}} + E^{(1)} + E^{(2)} + E^{(3)} + \dots \quad (6.130)$$

We will compute up to  $E^{(3)}$  here.

We obtain for the resolvent operator

$$\hat{R} = \sum_X \frac{1}{\Delta e_X} |\Phi_X\rangle \langle \Phi_X|, \quad \Delta e_X = \sum_{j=1}^{\#X} (\epsilon_{i_j} - \epsilon_{a_j}). \quad (6.131)$$

The first-order energy is

$$E^{(1)} = \langle \Phi | \{\hat{W}\} | \Phi \rangle = 0, \quad (6.132)$$

by Wick's Theorem.

The second-order energy is

$$E^{(2)} = \langle \Phi | \{\hat{W}\} \hat{R} \{\hat{W}\} | \Phi \rangle = \sum_X \frac{1}{\Delta e_X} |\langle \Phi | \{\hat{W}\} | \Phi_X \rangle|^2, \quad (6.133)$$

where only the doubles-excitations  $X = \binom{A \ B}{I \ J}$  will contribute. To see this, note that since  $\hat{W}$  is a two-body operator, at most two particle-hole pairs can be excited. Moreover,

$$\langle \Phi | \{\hat{W}\} | \Phi_i^a \rangle = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle \Phi | \{c_p^\dagger c_q^\dagger c_s c_r\} \{b_a^\dagger b_i^\dagger\} | \Phi \rangle = 0, \quad (6.134)$$

since there is no way to generate a nonzero fully contracted term.

Thus we are left with doubles only:

$$\begin{aligned}\langle \Phi | \{\hat{W}\} | \Phi_{ij}^{ab} \rangle &= \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle \Phi | \{c_p^\dagger c_q^\dagger c_s c_r\} \{b_a^\dagger b_b^\dagger b_j^\dagger b_i^\dagger\} | \Phi \rangle \\ &= \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} (\delta_{rs} \delta_{sb} - \delta_{rb} \delta_{sa}) (\delta_{qj} \delta_{pi} - \delta_{pj} \delta_{qi}) = w_{ab}^{ij}.\end{aligned}\quad (6.135)$$

We used that  $c_s$  and  $c_r$  must be contracted with  $b_a^\dagger$  or  $b_b^\dagger$ , and that  $c_p^\dagger$  or  $c_q^\dagger$  must be contracted with  $b_i^\dagger$  or  $b_j^\dagger$ . This leaves only four contributions, whose signs can be conveniently factored as above.

We assemble the second-order contribution:

$$E^{(2)} = \sum_{i < j} \sum_{a < b} \frac{|w_{ab}^{ij}|^2}{\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b}.\quad (6.136)$$

The factor 1/4 is eliminated by restricting the summation region.

The third-order energy  $E^{(3)}$  can be computed in a similar fashion,

$$\begin{aligned}E^{(3)} &= \langle \Phi | \{\hat{W}\} \hat{R} \{\hat{W}\} \hat{R} \{\hat{W}\} | \Phi \rangle - E^{(1)} \langle \Phi | \{\hat{W}\} \hat{R}^2 \{\hat{W}\} | \Phi \rangle \\ &= \sum_{XY} \frac{1}{\Delta e_X \Delta e_Y} \langle \Phi | \{\hat{W}\} | \Phi_X \rangle \langle \Phi_X | \{\hat{W}\} | \Phi_Y \rangle \langle \Phi_Y | \{\hat{W}\} | \Phi \rangle.\end{aligned}\quad (6.137)$$

where we used  $E^{(1)} = 0$ . The sum is only over  $X$  and  $Y$  being double excitations. We need to compute the middle factor on the last line, where  $X = (ij, ab)$  and  $Y = (kl, cd)$  are double excitations.

$$\begin{aligned}\langle \Phi_{ij}^{ab} | \{\hat{W}\} | \Phi_{kl}^{cd} \rangle &= \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle \Phi | \{b_i b_j b_b b_a\} \{c_p^\dagger c_q^\dagger c_s c_r\} \{b_c^\dagger b_d^\dagger b_l^\dagger b_k^\dagger\} | \Phi \rangle + 7 \text{ similar} \\ &\quad + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle \Phi | \{b_i b_j b_b b_a\} \{c_p^\dagger c_q^\dagger c_s c_r\} \{b_c^\dagger b_d^\dagger b_l^\dagger b_k^\dagger\} | \Phi \rangle + 7 \text{ similar} \\ &\quad + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \langle \Phi | \{b_i b_j b_b b_a\} \{c_p^\dagger c_q^\dagger c_s c_r\} \{b_c^\dagger b_d^\dagger b_l^\dagger b_k^\dagger\} | \Phi \rangle \\ &= \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} (\delta_{rc} \delta_{sd} - \delta_{rd} \delta_{cs}) (\delta_{ap} \delta_{bq} - \delta_{aq} \delta_{bp}) (\delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}) \\ &\quad + \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} (\delta_{ir} \delta_{js} - \delta_{is} \delta_{jr}) (\delta_{bd} \delta_{ac} - \delta_{bc} \delta_{ad}) (\delta_{pk} \delta_{ql} - \delta_{pl} \delta_{qk}) \\ &\quad + P(a, b) P(i, j) P(c, d) P(l, k) \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} (\delta_{ap} \delta_{ql} - \delta_{qa} \delta_{pl}) (\delta_{js} \delta_{rc} - \delta_{jr} \delta_{sc}) \delta_{ik} \delta_{bd} \\ &= w_{cd}^{ab} (\delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}) + w_{ij}^{kl} (\delta_{bd} \delta_{ac} - \delta_{bc} \delta_{ad}) + P(a, b) P(c, d) P(i, j) P(k, l) w_{cj}^{al} \delta_{ik} \delta_{bd} \\ &\quad (6.138)\end{aligned}$$

We have drawn 3 out of 32 possible fully contracted terms, falling into “natural groups” of 8 + 8 + 16. The rest of the fully contracted terms are obtained by permuting pairs of contractions, yielding minus signs,

as indicated. For example, in the first set of 8 contractions, we permute  $ba$ ,  $cd$ , and  $ij$ . The “groups” are defined as follows: (1)  $c_p^\dagger c_q^\dagger$  is contracted to the left, and  $c_s c_r$  to the right, (2) vice versa, and (3),  $c_p^\dagger c_s$  is contracted to the left and  $c_q^\dagger c_r$  to the right, and vice versa.

For the last 16 terms, we have opted to introduce the permutation symbol  $P(x, y)$  which is defined by

$$P(x, y)f(x, y, \dots) = f(x, y, \dots) - f(y, x, \dots).$$

We obtain, noting that the 16 terms from the last group of permutations will give identical results (exercise),

$$\begin{aligned} E^{(3)} &= \frac{1}{16} \sum_{ab} \sum_{cd} \sum_{ij} \sum_{kl} \frac{1}{\Delta e_{ij}^{ab} \Delta e_{kl}^{cd}} w_{ab}^{ij} w_{cd}^{ab} w_{kl}^{cd} (\delta_{ik} \delta_{jl} - \delta_{il} \delta_{jk}) \\ &\quad + \frac{1}{16} \sum_{ab} \sum_{cd} \sum_{ij} \sum_{kl} \frac{1}{\Delta e_{ij}^{ab} \Delta e_{kl}^{cd}} w_{ab}^{ij} w_{ij}^{kl} w_{kl}^{cd} (\delta_{bd} \delta_{ac} - \delta_{bc} \delta_{ad}) \\ &\quad + \frac{1}{16} \sum_{ab} \sum_c \sum_{ij} \sum_l \frac{1}{\Delta e_{ij}^{ab} \Delta e_{il}^{cb}} P(a, b) P(c, b) P(i, j) P(i, l) w_{ab}^{ij} w_{cj}^{al} w_{il}^{cb} \\ &= \frac{1}{8} \sum_{ab} \sum_{cd} \sum_{ij} \frac{w_{ab}^{ij} w_{cd}^{ab} w_{ij}^{cd}}{(\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)(\epsilon_c + \epsilon_d - \epsilon_i - \epsilon_j)} \\ &\quad + \frac{1}{8} \sum_{ab} \sum_{ij} \sum_{kl} \frac{w_{ab}^{ij} w_{ij}^{kl} w_{kl}^{ab}}{(\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)(\epsilon_a + \epsilon_b - \epsilon_k - \epsilon_l)} \\ &\quad + \sum_{abc} \sum_{ijk} \frac{w_{ab}^{ij} w_{ik}^{cb} w_{cj}^{ak}}{(\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)(\epsilon_b + \epsilon_c - \epsilon_i - \epsilon_k)} \end{aligned} \tag{6.139}$$

Looking at the energy correction expression for MP2 we find a cost (in the sense of the number of operations needed) of the order  $O(N^2 N_{\text{vir}}^2)$ , where  $N_{\text{vir}} + N$  is the total number of single-particle functions. Similarly, the cost of MP3 becomes  $O(N^2 N_{\text{vir}}^4) + O(N^4 N_{\text{vir}}^2) + O(N^3 N_{\text{vir}}^3)$ . Already at MP3, the cost is quite high. In quantum chemistry, the gold standard is CCSD(T) [see chapter on coupled-cluster theory], which has a cost of  $O((N + N_{\text{vir}})^6)$ . This cost is comparable to MP3. Thus, considering the convergence problems of RSPT expansions in general, it may be argued that in most circumstances, higher-order MP expansions are not that useful.

### 6.4.6 Exercises

**Exercise 6.7.** (Basic RSPT.) Prove Eq. (6.86),

$$|\Psi^{(2)}\rangle = \hat{R}[\hat{V} - \langle\Phi|\hat{V}|\Phi\rangle]\hat{R}\hat{V}|\Phi\rangle$$

△

**Exercise 6.8.** (Basic RSPT.) Derive the fourth and fifth order perturbation theory corrections to the energy in general RSPR, from Theorem 6.4. Next, verify that the bracketing technique gives the correct answer for these energy corrections.

△

**Exercise 6.9.** (The two-state model.) Prove that  $\hat{R}\hat{V}\hat{R} = 0$ . Compute  $E^{(n)}$ ,  $n = 4, 5, 6$ , for the two-state model, continuing the calculations in the section on the two-state model. Use the bracketing technique to derive the terms. Verify that your calculations match the terms in the Taylor expansion 6.99.  $\triangle$

**Exercise 6.10.** (A three-state model.)

a) Let  $H = H_0 + \lambda V$ , with

$$H_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}, \quad V = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 2 & i \\ 0 & -i & 0 \end{pmatrix}.$$

Compute RSPT for the ground-state energy up to fifth order, using the bracketing technique if you like.

Using your favorite computing environment, make a plot of the three eigenvalues as function of  $\lambda \in [-2, 2]$ . Can you estimate the radius of convergence of the RSPT series? Will the series converge for  $\lambda = 1$ ?

b) Write a program that numerically computes  $|\Psi^{(n)}\rangle$  and  $E^{(n)}$  using Theorem 6.4. Plot  $E^{(n)}$  and  $\|\Psi^{(n)}\|$  as functions of  $n$  in the same axis system. Use about 100 iterations. Experiment with different values of  $\lambda$  and discuss the behavior of each order's contributions for higher  $n$ . Does your radius of convergence from a) seem to be about right?

$\triangle$

**Exercise 6.11.** (MBPT.) In this exercise, you verify the formula for the zero-order energy levels in MBPT. Compute, for example by applying the Generalized Wick's Theorem to  $\hat{K}$ ,

$$\epsilon_X = \langle \Phi_X | \hat{K} | \Phi_X \rangle = \sum_i \kappa_i + \sum_{j=1}^{\#X} (\kappa_{a_j} - \kappa_{i_j}). \quad (6.140)$$

Here,  $X$  is a general excitation.

$\triangle$

**Exercise 6.12.** (MBPT.) Compute the following matrix elements. Use, e.g., elementary operator replacement considerations or the Generalized Wick's Theorem.

$$\begin{aligned} \langle \Phi_i^a | \hat{L}^{(1)} | \Phi \rangle &= \langle \phi_a | \hat{\ell}^{(1)} | \phi_i \rangle \\ \langle \Phi_i^a | \hat{L}^{(2)} | \Phi \rangle &= \sum_j \langle \phi_j \phi_a | \hat{\ell}^{(2)} | \phi_j \phi_i \rangle_{AS} \\ \langle \Phi_{ij}^{ab} | \hat{L}^{(2)} | \Phi \rangle &= \langle \phi_a \phi_b | \hat{\ell}^{(2)} | \phi_i \phi_j \rangle_{AS}. \end{aligned}$$

Show that there are no other nonzero matrix elements involving  $|\Phi\rangle$ .

$\triangle$



**Exercise 6.13.** (MBPT.) Write out  $E^{(3)}$  in the same fashion as Eq. (6.125).

△

**Exercise 6.14.** (Strenuous.) (MPPT.) Compute a closed formula for the MP4 correction to the energy. Follow the approach taken for MP3 above, using the bracketing technique and the generalized Wick's Theorem. You will need to compute matrix elements on the form

$$\langle \Phi_X | \{\hat{W}\} | \Phi_Y \rangle, \quad (6.141)$$

where  $X$  is a double excitation, and  $Y$  is a single, double, triple, or quadruple excitation.

△

# Chapter 7

## The Electron Gas

### 7.1 The Jellium Model

In this section we discuss the *jellium model*: an *infinite* system of interacting electrons and a uniform background charge, so that the system, on average, is neutral. This is also called *the electron gas*, and is a first-approximation to, among other things, a metal. A surprising amount of insight can be obtained from the jellium model, and we will here only scratch the surface.

The electron gas is an important theoretical model. It is useful as a model of metals, semiconductor heterostructures, etc., and it is the theoretical foundation of *density-functional theory* (DFT), a very popular computational technique in chemistry and solid-state physics.

From a many-body theoretical perspective, the electron gas is particularly interesting because it is an example of a system where the HF equations can be solved *analytically*. It also displays divergent terms in the RSPT series, another interesting phenomenon.

An infinite system is hard to treat mathematically. It is natural to start with a finite system and then take a limit afterwards. Throw  $N$  electrons in a box of sides  $L$  and volume  $\Omega = L^3$ . The average density is  $\bar{\rho} = N/\Omega$ , and we add a background charge  $eN$  to balance the electron charge  $-eN$ . Smearing the background charge uniformly gives a charge density  $e\bar{\rho}$  – hence a “jelly” in which the electrons move. Making the system electrically neutral is essential to avoid infinities: even a uniform distribution of charges has an infinite Coulomb interaction energy.

After having obtained the results for this box-truncated jellium, one then considers the *thermodynamic limit*, sending  $L \rightarrow +\infty$  and  $N \rightarrow +\infty$  together, such that  $\bar{\rho}$  is kept constant. Note that we measure physical parameters using *size-intensive* quantities in the thermodynamic limit.  $N$  and  $L^3$  are *size-extensive*, and not invariant with system size.

#### 7.1.1 Fourier series and plane-wave basis sets

For convenience, we impose *periodic boundary conditions* on our box. Any periodic function of reasonable behavior can be summed as a *Fourier series*,

$$f(\vec{r}) = \frac{1}{\Omega} \sum_{\vec{k}} \tilde{f}(\vec{k}) e^{i\vec{k} \cdot \vec{r}}, \quad (7.1)$$

where the sum extends over  $\vec{k}$  such that

$$\vec{k} = 2\pi\vec{\kappa}/L, \quad \kappa_x, \kappa_2, \kappa_3 \in \mathbb{Z}. \quad (7.2)$$

(It is an exercise to show that the Fourier modes  $\exp(i\vec{k} \cdot \vec{r})$  are then periodic functions.) We use the notation  $\mathcal{K}_L = \{\vec{k}\}$  for the set of wavenumbers.

The Fourier coefficients are given by

$$\tilde{f}(\vec{k}) = \int_B f(\vec{r}) e^{-i\vec{k} \cdot \vec{r}} d\vec{r}, \quad (7.3)$$

where the integral extends over the box  $B = [-L/2, +L/2]^3$ .

The kinetic energy operator of a single electron is

$$\hat{t} = -\frac{\hbar^2}{2m} \nabla^2, \quad (7.4)$$

whose eigenfunctions are the Fourier modes,

$$\varphi_{\vec{k}}(\vec{r}) = \frac{1}{\Omega^{1/2}} e^{i\vec{k} \cdot \vec{r}}. \quad (7.5)$$

These functions are orthonormal,

$$(\varphi_{\vec{k}} | \varphi_{\vec{k}'} ) = \delta_{\vec{k}, \vec{k}'}. \quad (7.6)$$

These are the common *plane-wave basis functions*. Note that the normalization factor differs from the Fourier series convention above.

The PW basis functions are very useful, since the kinetic energy is diagonal in this basis. Indeed, the momentum operator of a single electron is

$$\hat{\vec{p}} = -i\hbar \nabla, \quad (7.7)$$

and when acting on a Fourier mode,

$$\hat{\vec{p}} \varphi_{\vec{k}}(\vec{r}) = \hbar \vec{k} \varphi_{\vec{k}}(\vec{r}). \quad (7.8)$$

Thus,  $\hbar \vec{k} = \vec{p}_{\vec{k}}$  is an eigenvalue for the momentum operator, and the plane-wave basis is an eigenbasis.

Adding spin to the picture, we get spin-orbitals on the form  $\phi_{\vec{k}, \sigma}(\vec{r}, s) = \varphi_{\vec{k}}(\vec{r}) \chi_{\sigma}(s)$ , forming our single-particle basis, with the corresponding creation operators  $c_{\vec{k}, \sigma}^{\dagger}$ .

**Exercise 7.1.** Prove that the plane-wave basis functions  $\varphi_{\vec{k}}(\vec{r})$  form an orthonormal set if and only if the wave numbers are of the form of Eq. (7.2), i.e.,  $k_{\alpha} = 2\pi \kappa_{\alpha}/L$  for each space direction  $\alpha = x, y, z$ .  $\triangle$

**Exercise 7.2.** Let  $\vec{k}_i, i = 1, \dots, N$  be momentum vectors in  $\mathcal{K}_L$ . Let  $|\Phi\rangle = |(\vec{k}_1 \alpha_1) \cdots (\vec{k}_N \alpha_N)\rangle$ , a Slater determinant.

Explain why  $|\Phi\rangle$  is an eigenfunction of the total momentum operator  $\hat{\vec{P}} = \sum_i \hat{\vec{p}}(i)$  and compute its eigenvalue. Repeat for the total kinetic energy operator  $\hat{T} = \sum_i \hat{t}(i)$ .  $\triangle$

### 7.1.2 Non-interacting electrons in a box: Fermi gas

Assuming that the electrons do not interact, i.e., we set the charge  $e = 0$ , we obtain the simple Hamiltonian

$$\hat{H} = \hat{T} = \sum_{i=1}^N \hat{t}(i). \quad (7.9)$$

The plane waves are eigenfunctions of  $\hat{t}$ ,

$$\hat{t}\varphi_{\vec{k}}(\vec{r}) = \frac{\hbar^2 k^2}{2m} \varphi_{\vec{k}}(\vec{r}), \quad (7.10)$$

which gives the second-quantized kinetic energy

$$\hat{T} = \sum_{\vec{k} \in \mathcal{K}} \frac{\hbar^2 k^2}{2m} \sum_{\sigma} c_{\vec{k},\sigma}^{\dagger} c_{\vec{k},\sigma}. \quad (7.11)$$

The ground-state of this Hamiltonian is the Slater determinant  $|\Phi\rangle$  where the first  $N/2$  lowest-energy orbitals are doubly occupied. The energy becomes

$$E_0 = 2 \sum_{\vec{k} \in \mathcal{K}_{\text{occ}}} \frac{\hbar^2 k^2}{2m}, \quad (7.12)$$

where the summation extends over the occupied orbitals, denoted by the set  $\mathcal{K}_{\text{occ}} \subset \mathcal{K}$ .

Suppose we fix the average density  $\bar{\rho}$  and increase the box size  $L$ . Then,  $N = \bar{\rho}N$  grows, and the set of occupied states  $\mathcal{K}_{\text{occ}}$  has a very large number of elements. For large  $L$  and  $N$ ,  $\mathcal{K}_{\text{occ}}$  consists of those grid points in  $\mathcal{K}$  that are approximately inside a sphere, since the single-electron kinetic energy only depends on  $k = |\vec{k}|$ , and it is increasing in  $k$ . For a large number of particles, it is more convenient to use the *radius*  $k_F$  to measure the number of particles in terms of  $\bar{\rho}$  and  $\Omega$ .

We now let  $k_F$  be the radius of this sphere (the Fermi radius). To a good approximation, we should be able to find  $N$  as a function of  $k_F$  and  $L$ . The energy of the fastest electron (the Fermi energy) will be

$$\epsilon_F = \frac{\hbar^2 k_F^2}{2m}. \quad (7.13)$$

This gives for the number of electrons

$$N = 2 \sum_{\vec{k} \in U} = 2 \left( \frac{L}{2\pi} \right)^3 \sum_{2\pi\vec{k}/L \in \mathcal{K}_{\text{occ}}} \left( \frac{2\pi}{L} \right)^3 \approx 2 \left( \frac{L}{2\pi} \right)^3 \int_{|\vec{k}| < k_F} d\vec{k}, \quad (7.14)$$

where we have used the definition of the Riemann integral in the last equality:  $\vec{k}$  is a vector of integers, and for large  $L$  we see that  $\vec{k} = 2\pi\vec{k}/L$  lies on a grid with grid spacing  $2\pi/L$  in each spatial direction. The volume of each “cell” in  $\vec{k}$ -space is  $(2\pi/L)^3$ . The error in the integral approximation is of order  $L^{-3}$ .

The integral computes the volume of the sphere  $|\vec{k}| < k_F$ . This gives

$$N \approx 2 \left( \frac{L}{2\pi} \right)^3 \frac{4\pi}{3} k_F^3. \quad (7.15)$$

We observe that  $N$  becomes proportional to  $\Omega = L^3$ . Dividing out,

$$\bar{\rho} = \frac{1}{3\pi^2} k_F^3. \quad (7.16)$$

The error in this last equality is  $O(L^{-6})$ , which we can safely ignore. We note that  $k_F$  can be used as a variable to describe the non-interacting gas, equivalent to  $\bar{\rho}$  via Eq. (7.16),

$$k_F = (3\pi^2 \bar{\rho})^{1/3}. \quad (7.17)$$

We can also express the density in terms of the Fermi energy,

$$\bar{\rho} = \frac{(2m)^{3/2}}{3\pi^2 \hbar^3} \epsilon_F^{3/2}. \quad (7.18)$$

We observe that the integral approximation argument is valid in more general terms: suppose we are given a subset  $V \subset \mathcal{K}$  and want to compute

$$S = \sum_{\vec{k} \in V} f(\vec{k}). \quad (7.19)$$

Repeating the above trick,

$$\frac{S}{L^3} = L^{-3} \left( \frac{L}{2\pi} \right)^3 \sum_{\vec{k} \in V} \left( \frac{2\pi}{L} \right)^3 f(\vec{k}) \approx (2\pi)^{-3} \int_V f(\vec{k}) d\vec{k}, \quad (7.20)$$

with error  $O(L^{-6})$ , a good approximation for large  $L$ .

We now compute the ground-state energy per volume unit, a size-intensive quantity:

$$\frac{E_0}{L^3} = \frac{2}{L^3} \sum_{|\vec{k}| < k_F} \frac{\hbar^2 k^2}{2m} \approx (2\pi)^{-3} \frac{\hbar^2}{m} \int_{|\vec{k}| < k_F} k^2 d\vec{k} = \frac{\hbar^2}{(2\pi)^3 m} 4\pi \int_0^{k_F} k^4 dk = \frac{2}{5} \frac{\hbar^2}{m} \frac{1}{(2\pi)^2} k_F^5. \quad (7.21)$$

The left-hand side is actually the energy *density*. Using Eq. (7.17) we obtain

$$\frac{E_0}{\Omega} = \frac{\hbar^2}{m} \frac{\pi^{4/3} 3^{5/3}}{10} \bar{\rho}^{5/3}. \quad (7.22)$$

Finally, we compute the energy per particle in terms of the Fermi energy,

$$\frac{E_0}{N} = \frac{E_0}{\Omega \bar{\rho}} = \frac{3}{5} \epsilon_F. \quad (7.23)$$

### 7.1.3 Interacting gas/Jellium: Hamiltonian

Gross/Runge/Heinonen Chapter 5 is useful here. Also, Ref. [8].

For interacting electrons, the Hamiltonian reads

$$\hat{H} = \hat{T} + \hat{W} + \hat{V}_{b-e} + \hat{V}_{b-b}, \quad (7.24)$$

where  $\hat{T}$  is the kinetic energy of the electrons, and  $\hat{W}$  is the inter-electronic repulsion, and where  $\hat{V}_{b-e}$  and  $\hat{V}_{b-b}$  are the potential energy terms from interactions between the background and the electrons and the background with itself, respectively.

We now compute the classical potential energies  $W$ ,  $V_{b-e}$ , and  $V_{b-b}$ , in order to better understand the quantized operators.

The electric potential energy between two charges  $Q$  and  $Q'$  at  $\vec{r}$  and  $\vec{r}'$ , respectively, is

$$V = \frac{1}{4\pi\epsilon_0} \frac{QQ'}{|\vec{r} - \vec{r}'|}. \quad (7.25)$$

Numerically, with  $e$  being the elementary electron charge,

$$\alpha \equiv \frac{e^2}{4\pi\epsilon_0} \approx 1.440 \text{ eV} \cdot \text{nm}. \quad (7.26)$$

We write

$$w(\vec{r}) = \frac{e^2}{4\pi\epsilon_0} \frac{1}{|\vec{r}|}, \quad (7.27)$$

so that  $V = QQ'e^{-2}w(\vec{r} - \vec{r}')$ . Typically,  $Q$  and  $Q'$  will be given in units of the elementary charge  $e$ , making this definition useful. Given two *charge densities in units of the elementary charge*  $e$ ,  $\rho_1(\vec{r})$  and  $\rho_2(\vec{r})$ , the total potential energy becomes

$$V_{12} = \iint \rho_1(\vec{r})\rho_2(\vec{s})w(\vec{r} - \vec{s}) d\vec{r}d\vec{s}. \quad (7.28)$$

The interaction of  $\rho_1(\vec{r})$  with itself (assuming indistinguishable particles) becomes

$$V_{11} = \frac{1}{2} \iint \rho_1(\vec{r})\rho_1(\vec{s})w(\vec{r} - \vec{s}) d\vec{r}d\vec{s}, \quad (7.29)$$

The factor  $1/2$  comes from the fact that the integral counts every interaction between two particles *twice* due to indistinguishability.

The charge density of  $N$  classical electrons is

$$e\rho_e(\vec{r}) = e \sum_{i=1}^N \delta(\vec{r} - \vec{r}_i). \quad (7.30)$$

Applying the formula for  $V_{11}$ ,

$$W = \frac{1}{2} \sum_{i \neq j}^N w(\vec{r}_i - \vec{r}_j), \quad (7.31)$$

where have to exclude  $i = j$  from the sum, since otherwise we get an infinite contribution for each electron interacting with itself.<sup>1</sup> We anticipate that for  $N \rightarrow +\infty$  this will give an infinite potential energy, since the Coulomb potential has infinite range. This is why we introduce the background charge. On the other hand, the background charge density is static and uniform over the box, with density  $\rho_b(\vec{r}) = -\bar{\rho}$ , so that the system in total is neutral. Thus, the potential energy  $V_{b-b}$  becomes

$$V_{b-b} = \frac{1}{2} \int_B \int_B \rho_b(\vec{r})\rho_b(\vec{s})w(\vec{r} - \vec{s}) d\vec{r}d\vec{s} = \frac{1}{2} \bar{\rho}^2 \int_B \int_B w(\vec{r} - \vec{s}) d\vec{r}d\vec{s}. \quad (7.32)$$

This is just a constant number, which is very large but finite, for finite  $L$ . However, we note that in the thermodynamic limit,  $V_{b-b}$  grows without bound and represents another singularity of the model.

Next, consider the potential  $V_{e-b}$ , which is the potential between a  $\delta$ -function sum and a constant density,

$$V_{e-b} = \iint \rho_b(\vec{s})\rho_e(\vec{r})w(\vec{r} - \vec{s}) d\vec{r}d\vec{s} = -\bar{\rho} \sum_{i=1}^N \int_B w(\vec{r}_i - \vec{s})^{-1} d\vec{s} \quad (7.33)$$

---

<sup>1</sup>Do electrons interact with themselves? Why should not an electron feel the potential set up by itself? Feynman spent large parts of his career addressing this question. He was awarded the Nobel Prize in Physics in 1965 (jointly with Schwinger and Tomonaga) for QED, and he spent some of his Nobel lecture talking about the electron's self-interaction, see [http://www.nobelprize.org/nobel\\_prizes/physics/laureates/1965/feynman-lecture.html](http://www.nobelprize.org/nobel_prizes/physics/laureates/1965/feynman-lecture.html).

Here, we had use for the formula for  $V_{12}$ . In a similar fashion as  $V_{b-b}$ , the integral on the right-hand side is finite, but very large and *negative*. In the limit of a large box, the integral blows up (or down, rather).

The problem with the infinities comes from the fact that the Coulomb interaction has infinite range, in the sense of scattering theory. We therefore introduce the *Yukawa potential* as a regularization

$$w(\vec{r}; \mu) = e^{-\mu|\vec{r}|} w(\vec{r}) = \frac{e^2}{4\pi\epsilon_0} e^{-\mu r} r^{-1}, \quad \mu > 0, \quad r = |\vec{r}|, \quad (7.34)$$

which has finite scattering range  $\sim \mu^{-1}$  and gives the Coulomb potential in the limit  $\mu \rightarrow 0$ . The idea is to use the Yukawa potential for a finite  $L$ , and see that *the nasty terms cancel*. Then we can take the thermodynamic limit, and observing that our results are have a well-defined limit as  $\mu \rightarrow 0$ .

The Fourier transform of the Yukawa potential is

$$\tilde{w}(\vec{q}; \mu) \equiv \int_B w(\vec{r}; \mu) e^{-i\vec{q} \cdot \vec{r}}, \quad (7.35)$$

so that

$$w(\vec{r}; \mu) = \frac{1}{\Omega} \sum_{\vec{q} \in \mathcal{K}} \tilde{w}(\vec{q}; \mu) e^{i\vec{q} \cdot \vec{r}}, \quad \vec{r} \in B. \quad (7.36)$$

If we assume that the box is large enough, this integral becomes, to an exponentially good approximation,

$$\tilde{w}(\vec{q}; \mu) \approx \frac{e^2}{4\pi\epsilon_0} \frac{4\pi}{\mu^2 + q^2}. \quad (7.37)$$

This is an exercise.

We rewrite the Hamiltonian in terms of the Fourier transform (7.36). First, we define

$$\hat{n}_{\vec{q}} \equiv \sum_{i=1}^N e^{-i\vec{q} \cdot \vec{r}_i}, \quad (7.38)$$

a one-body operator. This gives

$$\begin{aligned} \hat{W} &= \frac{1}{2} \sum_{i \neq j} w(\vec{r}_i - \vec{r}_j; \mu) = \frac{1}{2\Omega} \sum_{i \neq j} \tilde{w}(\vec{q}; \mu) e^{i\vec{q} \cdot (\vec{r}_i - \vec{r}_j)} \\ &= \frac{1}{2\Omega} \sum_{\vec{q} \in \mathcal{K}} \tilde{w}(\vec{q}; \mu) (\hat{n}_{-\vec{q}} \hat{n}_{\vec{q}} - \hat{N}), \end{aligned} \quad (7.39)$$

where the number operator compensates for including  $i = j$  in the sum.

Next, we consider  $\hat{V}_{b-b}$ : It can be shown, that to a good approximation, and for a large enough box,

$$\int_B \int_B w(\vec{r} - \vec{s}) d\vec{r} d\vec{s} \approx \Omega \tilde{w}(\vec{0}, \mu). \quad (7.40)$$

This gives

$$\hat{V}_{b-b} = \frac{1}{2} \bar{\rho}^2 \Omega \tilde{w}(\vec{0}, \mu). \quad (7.41)$$

Similarly,

$$\hat{V}_{e-b} = -\bar{\rho} N \tilde{w}(\vec{0}, \mu). \quad (7.42)$$

We note that the divergent terms  $\hat{V}_{b-b}$  and  $\hat{V}_{e-b}$  are proportional to  $\tilde{w}(\vec{0}; \mu)$ . Thus, we separate out the  $\vec{q} = \vec{0}$  term of  $\hat{W}$  (noting that  $\hat{n}_0^2 = \hat{N}^2$ ),

$$\hat{H} = \hat{T} + \hat{W}_0 + \frac{1}{2\Omega} \tilde{w}(\vec{0}, \mu) (N^2 - N) - \bar{\rho} N \tilde{w}(\vec{0}, \mu) + \frac{1}{2} \bar{\rho}^2 \Omega \tilde{w}(\vec{0}, \mu) \quad (7.43)$$

We note that the divergencies are canceled, since

$$\frac{1}{2\Omega}(N^2 - N) - \bar{\rho}N + \frac{1}{2}\bar{\rho}^2\Omega = \frac{1}{2\Omega}(N^2 - N) - \frac{1}{\Omega}N^2 + \frac{1}{2\Omega}N^2 = -\frac{1}{2}\bar{\rho}. \quad (7.44)$$

which is a constant number. As we compute the energy *density* in the end, the contribution to the energy becomes  $-\frac{1}{2}\bar{\rho}/\Omega$  which can be safely ignored.

Thus, the Hamiltonian becomes, for fixed  $\bar{\rho}$  and large boxes,

$$\hat{H} = \hat{T} + \hat{W}_0, \quad \hat{W}_0 = \frac{1}{2\Omega} \sum_{\vec{q} \neq 0} \tilde{w}(\vec{q}; \mu) (\hat{n}_{-\vec{q}} \hat{n}_{\vec{q}} - \hat{N}). \quad (7.45)$$

#### 7.1.4 Hamiltonian in second quantization

We now express the Hamiltonian in second quantization using the plane-wave basis. To this end, first we consider the matrix elements of  $\hat{n}_{\vec{q}} = \sum_i e^{-i\vec{q} \cdot \vec{r}_i}$ , which is instructive:

$$(\varphi_{\vec{k}} | e^{-i\vec{q} \cdot \vec{r}} | \varphi_{\vec{\ell}}) = \frac{1}{\Omega} \int e^{-i\vec{k} \cdot \vec{r}} e^{-i\vec{q} \cdot \vec{r}} e^{+i\vec{\ell} \cdot \vec{r}} d\vec{r} = \frac{1}{\Omega} \int e^{i(-\vec{k}-\vec{q}+\vec{\ell}) \cdot \vec{r}} d\vec{r} = \delta_{-\vec{k}-\vec{q}+\vec{\ell}, \vec{0}}, \quad (7.46)$$

by the orthonormality of the  $\varphi_{\vec{k}}$ , for  $\vec{k} \in \mathcal{K}$ .

Summing up the operator, we obtain

$$\hat{n}_{\vec{q}} = \sum_{\vec{k}, \vec{\ell}} \sum_{\alpha} (\varphi_{\vec{k}} | e^{-i\vec{q} \cdot \vec{r}} | \varphi_{\vec{\ell}}) c_{\vec{k}\alpha}^{\dagger} c_{\vec{\ell}\alpha} = \sum_{\vec{k}, \alpha} c_{\vec{k}, \alpha}^{\dagger} c_{\vec{k}+\vec{q}, \alpha}, \quad (7.47)$$

where we used that  $\hat{n}_{\vec{q}}$  does not depend on spin. Thus,  $\hat{n}_{\vec{q}}$  is a *shift operator*, that annihilates a particle with wavenumber  $\vec{\ell}$  and inserts one with wavenumber  $\vec{\ell} - \vec{q}$ .

We note that

$$\hat{n}_{\vec{q}}^{\dagger} = \hat{n}_{-\vec{q}}. \quad (7.48)$$

Using the fundamental anticommutator and Eq. (7.47), it is straightforward to show that

$$\hat{n}_{-\vec{q}} \hat{n}_{\vec{q}} = \hat{N} + \sum_{\vec{k}, \vec{\ell}} \sum_{\alpha, \beta} c_{\vec{k}, \alpha}^{\dagger} c_{\vec{\ell}, \beta}^{\dagger} c_{\vec{\ell}+\vec{q}, \beta} c_{\vec{k}-\vec{q}, \alpha}. \quad (7.49)$$

We note that  $\hat{n}_{-\vec{q}} \hat{n}_{\vec{q}}$  conserves total momentum  $\hbar k \vec{K} = \hbar k \sum_{i=1}^N \vec{k}_i$ .

Plugging into  $\hat{W}_0$  we obtain

$$\hat{W}_0 = \frac{1}{2\Omega} \sum_{\vec{q} \neq 0} \tilde{w}(\vec{q}; \mu) \sum_{\vec{k}, \vec{\ell}} \sum_{\alpha, \beta} c_{\vec{k}, \alpha}^{\dagger} c_{\vec{\ell}, \beta}^{\dagger} c_{\vec{\ell}+\vec{q}, \beta} c_{\vec{k}-\vec{q}, \alpha}. \quad (7.50)$$

A useful observation is that *the operator  $\hat{T} + \hat{W}_0$  conserves the total momentum  $\sum_i \vec{k}_i$* . The same is of course true for the kinetic energy operator, which is diagonal in the plane-wave basis.

We explicitly include  $\vec{q} = 0$  in the sum by defining

$$\tilde{w}_0(\vec{q}; \mu) = \tilde{w}(\vec{q}; \mu) (1 - \delta_{\vec{q}, \vec{0}}). \quad (7.51)$$

We write  $\hat{W}_0$  in the following way:

$$\hat{W}_0 = \frac{1}{2\Omega} \sum_{\vec{k}, \vec{\ell}} \sum_{\vec{k}', \vec{\ell}'} \delta_{\vec{k}+\vec{\ell}, \vec{k}'+\vec{\ell}'} \tilde{w}_0(\vec{\ell}' - \vec{\ell}) \sum_{\alpha, \beta} c_{\vec{k}, \alpha}^{\dagger} c_{\vec{\ell}, \beta}^{\dagger} c_{\vec{\ell}', \beta} c_{\vec{k}', \alpha}. \quad (7.52)$$

From this, we can read off the non-antisymmetrized matrix elements of  $\hat{W}_0$ ,

$$\langle \phi_{\vec{k}, \alpha} \phi_{\vec{\ell}, \beta} | \hat{W}_0 | \phi_{\vec{k}', \alpha'} \phi_{\vec{\ell}', \beta'} \rangle = \delta_{\alpha\alpha'} \delta_{\beta\beta'} \delta_{\vec{k}+\vec{\ell}, \vec{k}'+\vec{\ell}'} \tilde{w}(\vec{\ell}' - \vec{\ell}; \mu) (1 - \delta_{\vec{\ell}, \vec{\ell}'}). \quad (7.53)$$



### 7.1.5 Hartree–Fock treatment

We attack the electron gas using HF theory. We do a *restricted Hartree–Fock* approach, where we use spin-orbitals that factor in a space- and a spin part, just like our basis.

We will show the remarkable fact that the plane-wave basis is a solution to the canonical (restricted) Hartree–Fock equations. To show this, we demonstrate that the Fock operator  $\hat{F} = \hat{T} + \hat{V}^{\text{direct}} - \hat{V}^{\text{exchange}}$  is diagonal in this basis.

Thus, we are going to plug our plane-wave basis  $\phi_{\vec{k},\sigma}$  into the Fock operator and show that it becomes diagonal. The occupied wavenumbers become precisely  $\mathcal{K}_{\text{occ}}$  as before, and each is doubly occupied. We use the symbols  $\vec{i}$  and  $\vec{j}$ , for occupied wavenumbers, and  $\vec{a}$ ,  $\vec{b}$ , etc, for unoccupied wavenumbers. The symbols  $\vec{k}$ ,  $\vec{l}$ ,  $\vec{q}$  denote general wavenumbers.

We have at our disposal Eq. (7.53). We compute  $\hat{V}^{\text{direct}}$ , and we use the fact that  $\hat{W}_0$  conserves total momentum:

$$\begin{aligned}\hat{V}^{\text{direct}} &= \sum_{\vec{k}_1, \tau_1} \sum_{\vec{k}_2, \tau_2} \sum_{\vec{i}, \sigma} \langle \phi_{\vec{k}_1, \tau_1} \phi_{\vec{i}, \sigma} | \hat{W}_0 | \phi_{\vec{k}_2, \tau_2} \phi_{\vec{i}, \sigma} \rangle c_{\vec{k}_1, \tau_1}^\dagger c_{\vec{k}_2, \tau_2} \\ &= \sum_{\vec{i}, \sigma} \sum_{\vec{k}, \tau} \langle \phi_{\vec{k}, \tau} \phi_{\vec{i}, \sigma} | \hat{W}_0 | \phi_{\vec{k}, \tau} \phi_{\vec{i}, \sigma} \rangle c_{\vec{k}, \tau}^\dagger c_{\vec{k}, \tau} = 0.\end{aligned}\quad (7.54)$$

To see the last equality, note that since the bra and the ket state are identical, we are left with terms from (??) containing  $\tilde{w}_0(\vec{i} - \vec{i}) = \tilde{w}_0(\vec{k} - \vec{k}) = \tilde{w}_0(\vec{0})$  only, and these are identically zero.

We now turn to the exchange potential:

$$\begin{aligned}\hat{V}^{\text{exchange}} &= \sum_{\vec{k}_1, \tau_1} \sum_{\vec{k}_2, \tau_2} \sum_{\vec{i}, \sigma} \langle \phi_{\vec{k}_1, \tau_1} \phi_{\vec{i}, \sigma} | \hat{W}_0 | \phi_{\vec{i}, \sigma} \phi_{\vec{k}_2, \tau_2} \rangle c_{\vec{k}_1, \tau_1}^\dagger c_{\vec{k}_2, \tau_2} \\ &= \sum_{\vec{i}, \sigma} \sum_{\vec{k}} \langle \phi_{\vec{k}, \sigma} \phi_{\vec{i}, \sigma} | \hat{W}_0 | \phi_{\vec{i}, \sigma} \phi_{\vec{k}, \sigma} \rangle c_{\vec{k}, \sigma}^\dagger c_{\vec{k}, \sigma} \\ &= \sum_{\vec{k}} \left( \sum_{\vec{i}} \frac{1}{\Omega} \tilde{w}(\vec{k} - \vec{i}; \mu) (1 - \delta_{\vec{k}, \vec{i}}) \right) \sum_{\sigma} c_{\vec{k}, \sigma}^\dagger c_{\vec{k}, \sigma}\end{aligned}\quad (7.55)$$

Unlike the direct part, the exchange operator does not vanish. On the other hand, it is manifestly diagonal in the plane-wave basis. We obtain for the Fock operator

$$\hat{F} = \hat{T} - \hat{V}^{\text{exchange}} = \sum_{\vec{k}} \left[ \frac{\hbar^2 k^2}{2m} - \left( \sum_{\vec{i}} \frac{1}{\Omega} \tilde{w}(\vec{k} - \vec{i}; \mu) (1 - \delta_{\vec{k}, \vec{i}}) \right) \right] \sum_{\alpha} c_{\vec{k}, \alpha}^\dagger c_{\vec{k}, \alpha} \equiv \sum_{\vec{k}} \epsilon_{\vec{k}} \sum_{\alpha} c_{\vec{k}, \alpha}^\dagger c_{\vec{k}, \alpha}. \quad (7.56)$$

The diagonal elements  $\epsilon_{\vec{k}}$  are therefore the canonical HF energies, which are doubly degenerate due to spin, and the spin-orbitals  $\phi_{\vec{k}, \alpha}$  are the canonical HF functions.

The HF energy depends only on  $k = |\vec{k}|$ , and we have:

$$\epsilon_{\vec{k}} \equiv \epsilon_k = \frac{\hbar^2 k^2}{2m} - \frac{1}{\Omega} \sum_{\vec{i}} \tilde{w}(\vec{k} - \vec{i}; \mu) (1 - \delta_{\vec{k}, \vec{i}}). \quad (7.57)$$

The HF energy is

$$E_{\text{HF}} = \langle \Phi | \hat{T} + \hat{W}_0 | \Phi \rangle = \sum_{\vec{i}, \alpha} t_{\vec{i}} - \frac{1}{2} \sum_{\vec{i}, \vec{j}, \alpha} \langle \phi_{\vec{j}, \alpha} \phi_{\vec{i}, \alpha} | \hat{W}_0 | \phi_{\vec{i}, \alpha} \phi_{\vec{j}, \alpha} \rangle \equiv E_K - E_{\text{exchange}} \quad (7.58)$$

where we used that the direct potential is identically zero, and only the exchange parts of the interaction matrix elements contribute.

Note that we have not *used* that the occupied wavenumbers are  $\mathcal{K}_{\text{occ}}$ . We could choose any set of  $N/2$  indices, and any such choice gives a solution to the canonical HF equations. It is natural to expect that the minimum HF energy is obtained by choosing precisely the Fermi sphere as occupied states, but it is not trivial. From now on we use the Fermi sphere, i.e.,  $\vec{i} \in \mathcal{K}_{\text{occ}}$  by definition.

### 7.1.6 Evaluation of sums in HF energies

We now evaluate the HF eigenvalue  $\epsilon_k$  and the total HF energy  $E_{\text{HF}}$  in the thermodynamic limit. We consider first the sum

$$S_k \equiv \frac{1}{\Omega} \sum_{\vec{i}} \tilde{w}(\vec{k} - \vec{i}; \mu) (1 - \delta_{\vec{k}, \vec{i}}), \quad (7.59)$$

so that  $\epsilon_k = \hbar^2 k^2 / 2m - S_k$ . We evaluate the sum as an integral, and set  $\mu = 0$  since the integral converges also for this limit:

$$S_k = \frac{4\pi\alpha}{(2\pi)^3} \int_{|\vec{s}| < k_F} \frac{1}{|\vec{s} - \vec{k}|^2} d\vec{s}. \quad (7.60)$$

Here,  $\alpha = e^2 / (4\pi\epsilon_0)$ , introduced to avoid the possible confusion of the Coulomb force constant  $k$  with the wavenumber  $k = |\vec{k}|$ . To evaluate the integral, we choose the  $z$ -axis in  $\vec{s}$ -space along  $\vec{k}$ , introduce spherical coordinates, and get

$$|\vec{s} - \vec{k}|^2 = s^2 + k^2 - 2ks \cos(\theta). \quad (7.61)$$

Inserting into the integral, we get

$$S_k = \frac{4\pi\alpha}{(2\pi)^3} 2\pi \int_{-\pi/2}^{\pi/2} \sin^2(\theta) d\theta \int_0^{k_F} s^2 ds [s^2 + k^2 - 2ks \cos(\theta)]^{-1}. \quad (7.62)$$

Introducing  $x = \cos(\theta)$  as integration variable, we can complete the calculation and obtain

$$S_k = \frac{\alpha}{\pi} \left[ k_F + \frac{1}{2k} (k_F^2 - k^2) \ln \left| \frac{k_F + k}{k_F - k} \right| \right]. \quad (7.63)$$

**Exercise 7.3.** Fill in the details of the above integration. △

Thus, we obtain

$$\epsilon_k = \frac{\hbar^2 k^2}{2m} - \frac{\alpha}{\pi} \left[ k_F + \frac{1}{2k} (k_F^2 - k^2) \ln \left| \frac{k_F + k}{k_F - k} \right| \right] \quad (7.64)$$

We now turn to the calculation of  $E_{\text{HF}}$ . First, we note that the kinetic energy  $E_K$  was calculated in the section about the noninteracting gas. There,  $k_F$  was expressed in terms of the density  $\bar{\rho}$  and vice versa,

$$\bar{\rho} = \frac{1}{3\pi^2} k_F^3. \quad (7.65)$$

The density is the same in the HF model, since the state  $|\Phi\rangle$  is the same. The kinetic energy in terms of  $k_F$  becomes

$$E_K = \Omega \frac{\hbar^2}{m} \frac{\pi^{4/3} 3^{5/3}}{10} [(3\pi^2) k_F^3]^{5/3} = \Omega \frac{\hbar^2}{m} \frac{1}{10\pi^2} k_F^5. \quad (7.66)$$

Now to the exchange energy.

$$\begin{aligned}
E_{\text{exchange}} &= \frac{1}{2} \sum_{\vec{i}\vec{j}\alpha} \langle \phi_{\vec{i}\alpha} \phi_{\vec{j}\alpha} | \hat{W}_0 | \phi_{\vec{j}\alpha} \phi_{\vec{i}\alpha} \rangle = \frac{1}{\Omega} \sum_{\vec{i}\vec{j}} \tilde{w}_0(\vec{i} - \vec{j}) = \sum_{\vec{j}} S_j \\
&\approx \frac{\Omega}{(2\pi)^3 \alpha} \int_{|\vec{j}| < k_F} S_j = \frac{\Omega}{(2\pi)^3 \alpha} 4\pi \int_0^{k_F} k^2 S_k dk.
\end{aligned} \tag{7.67}$$

This integral can be carried out by elementary means, to give

$$E_{\text{exchange}} = \Omega \frac{2\alpha}{(2\pi)^3} k_F^4. \tag{7.68}$$

In total, therefore, we get the energy density

$$\begin{aligned}
\frac{E_{\text{HF}}}{\Omega} &= \frac{\hbar^2}{m} \frac{1}{10\pi^2} k_F^5 - \frac{2\alpha}{(2\pi)^3} k_F^4 \\
&= \frac{\hbar^2}{m} \frac{3^{5/3} \pi^{4/3}}{10} \bar{\rho}^{5/3} - \alpha \frac{3^{4/3}}{4\pi^{1/3}} \bar{\rho}^{4/3}.
\end{aligned} \tag{7.69}$$

valid in the limit  $\Omega \rightarrow +\infty$ .

**Exercise 7.4.** Fill in the details: Compute the integral in Eq. (7.67) to obtain Eq. (7.68) △

**Exercise 7.5.** a) Show that the HF energy density (7.69) as a function of the average electron density  $\bar{\rho}$  can be written

$$\frac{E_{\text{HF}}}{\Omega} = \frac{\hbar^2}{m} \frac{3^{5/3} \pi^{4/3}}{10} \bar{\rho}^{5/3} - 2\alpha \frac{3^{4/3}}{8\pi^{1/3}} \bar{\rho}^{4/3}. \tag{7.70}$$

b) Show that the HF energy per particle can be written

$$\frac{E_{\text{HF}}}{N} = \frac{\hbar^2}{m} \frac{3^{5/3} \pi^{4/3}}{10} \bar{\rho}^{2/3} - 2\alpha \frac{3^{4/3}}{8\pi^{1/3}} \bar{\rho}^{1/3}. \tag{7.71}$$

△

### 7.1.7 Spin polarization: Unrestricted Hartree–Fock

We have derived the HF energy under the assumption that the number of electrons with spins up and down were equal. But could it be that we actually may lower the energy by polarizing the gas?

As argued, *any* set of occupied wavenumbers will lead to a diagonal Fock operator. Thus, we may set  $N = N_{\uparrow} + N_{\downarrow}$ , and consider the *unrestricted* HF ansatz instead of the restricted ansatz. Since  $\hat{V}^{\text{HF}}$  does not couple different spins, we obtain the UHF energy as the sum of two parts, coming from the two Fermi spheres of different spins. Setting  $\bar{\rho} = \rho_{\uparrow} + \rho_{\downarrow}$ , we get

$$E_{\text{UHF}} = \frac{\hbar^2}{m} \frac{3^{5/3} \pi^{4/3}}{10} (\rho_{\uparrow}^{5/3} + \rho_{\downarrow}^{5/3}) - \alpha \frac{3^{4/3}}{4\pi^{1/3}} (\rho_{\uparrow}^{4/3} + \rho_{\downarrow}^{4/3}). \tag{7.72}$$

We may try to minimize this energy under the constraint that  $\bar{\rho}$  is fixed. Introduce the *fractional polarization*  $p = (N/2 - N_\uparrow)/(N/2) = 1 - 2\rho_\uparrow/\bar{\rho}$ , so that  $\rho_\uparrow = (1 - p)\bar{\rho}/2$  and  $\rho_\downarrow = (1 + p)\bar{\rho}/2$ , and write

$$E_{\text{UHF}} = A((1 - p)^{5/3} + (1 + p)^{5/3}) - B((1 - p)^{4/3} + (1 + p)^{4/3}), \quad (7.73)$$

with

$$A = \frac{\hbar^2 (\frac{3}{2})^{5/3} \pi^{4/3}}{m \cdot 10} \bar{\rho}^{5/3}, \quad B = \alpha \frac{(\frac{3}{2})^{4/3}}{4\pi^{1/3}} \bar{\rho}^{4/3} \quad (7.74)$$

Minimizing  $E_{\text{UHF}}$  with respect to  $0 \leq p \leq 1$ , one finds that the partially polarized state ( $0 < p < 1$ ) is never a minimum, but that  $p = 0$  or  $p = 1$  is a minimum, depending on the relative magnitude of the constants  $A$  and  $B$ , i.e., on the density  $\bar{\rho}$ . There is a critical density  $\bar{\rho}_B$  where the model undergoes a so-called Bloch transition, where  $p = 0$  (paramagnetic) jumps to  $p = 1$  (ferromagnetic, all spins aligned).

### 7.1.8 Perturbation theory for jellium is divergent

In this section, we demonstrate that the second-order energy correction of Rayleigh–Schrödinger perturbation theory is divergent in itself. Thus, it is not the *series* as such, but the individual terms in the series that are problematic in PT for jellium.

We treat  $\hat{W}_0$  as a perturbation of  $\hat{T}$ , i.e., in the terminology of Sec. 6.4,  $\hat{K} = \hat{T}$  and  $\hat{L} = \hat{W}_0$ . (We could, alternatively, use  $\hat{F} = \hat{T} - \hat{V}_{\text{exchange}}$  as the unperturbed operator, and  $\hat{W}_0 + \hat{V}_{\text{exchange}}$  as the perturbation. This leads to a different formula for the second-order energy, which we do not analyze here.)

The treatment follows Raimis Ch. 3.6 closely, but with our own notation.

The zeroth order state  $|\Phi\rangle$  is the ground-state of  $\hat{T}$ , the state where all  $(\vec{k}, \alpha)$  are occupied with  $|\vec{k}| < k_F$ . In the thermodynamic limit, we have

$$\frac{E^{(0)}}{\Omega} = \frac{E_K}{\Omega}, \quad (7.75)$$

as computed previously.

The first-order energy is

$$\frac{E^{(1)}}{\Omega} = \frac{1}{\Omega} \langle \Phi | \hat{W}_0 | \Phi \rangle = -\frac{1}{\Omega} E_{\text{exchange}}. \quad (7.76)$$

Here, we used that  $|\Phi\rangle$  is actually the HF state. Thus,

$$\frac{1}{\Omega} (E^{(0)} + E^{(1)}) = \frac{1}{\Omega} E_{\text{HF}}. \quad (7.77)$$

Now to the second-order energy correction. Since  $|\Phi\rangle$  is the HF state, Brillouin's Theorem gives that the second-order energy correction does not have contributions from one-particle-one-hole Slater determinants. We are left with

$$\frac{1}{\Omega} E^{(2)} = \frac{1}{4\Omega} \sum_{IJAB} (t_I - t_A + t_J - t_B)^{-1} |\langle \phi_A \phi_B | \hat{W}_0 | \phi_I \phi_J \rangle_{\text{AS}}|^2. \quad (7.78)$$

Here,

$$I = (\vec{i}, \alpha), \quad J = (\vec{j}, \alpha'), \quad A = (\vec{a}, \beta), \quad B = (\vec{a}, \beta'). \quad (7.79)$$

$$t_K = \frac{\hbar^2}{2m} |\vec{k}|^2. \quad (7.80)$$

We have

$$\langle \phi_I \phi_J | \hat{W}_0 | \phi_A \phi_B \rangle_{\text{AS}} = \langle \phi_{\vec{i}, \alpha} \phi_{\vec{j}, \alpha'} | \hat{W}_0 | \phi_{\vec{a}, \beta} \phi_{\vec{a}, \beta'} \rangle - \langle \phi_{\vec{i}, \alpha} \phi_{\vec{j}, \alpha'} | \hat{W}_0 | \phi_{\vec{b}, \beta'} \phi_{\vec{a}, \beta} \rangle, \quad (7.81)$$

where

$$\langle \phi_{\vec{k}_1, \alpha_1} \phi_{\vec{k}_2, \alpha_2} | \hat{w}_0 | \phi_{\vec{\ell}_1, \beta_1} \phi_{\vec{\ell}_2, \beta_2} \rangle = \frac{1}{\Omega} \delta_{\alpha_1 \beta_1} \delta_{\alpha_2 \beta_2} \delta_{\vec{k}_1 + \vec{k}_2, \vec{\ell}_1 + \vec{\ell}_2} \tilde{w}_0(\vec{k}_1 - \vec{\ell}_1; \mu). \quad (7.82)$$

Thus,

$$\begin{aligned} \langle \phi_I \phi_J | \hat{w}_0 | \phi_A \phi_B \rangle_{AS} &= \frac{1}{\Omega} \delta_{\vec{i} + \vec{j}, \vec{a} + \vec{b}} \left[ \delta_{\alpha \beta} \delta_{\alpha' \beta'} \tilde{w}_0(\vec{i} - \vec{a}; \mu) - \delta_{\alpha \beta'} \delta_{\beta \alpha'} \tilde{w}_0(\vec{i} - \vec{b}; \mu) \right] \\ &= \frac{4\pi\alpha}{\Omega} \delta_{\vec{i} + \vec{j}, \vec{a} + \vec{b}} \left[ \delta_{\alpha \beta} \delta_{\alpha' \beta'} (\mu^2 + |\vec{i} - \vec{a}|^2)^{-1} - \delta_{\alpha \beta'} \delta_{\beta \alpha'} (\mu^2 + |\vec{i} - \vec{b}|^2)^{-1} \right] \end{aligned} \quad (7.83)$$

We are going to sum over  $a$ ,  $b$ ,  $i$ , and  $j$ , and we note that total spin projection must be conserved,  $\alpha + \alpha' = \beta + \beta'$ . For a simpler integral analysis later on, we split the contributions into the cases  $\alpha = -\alpha'$  (anti-parallel spins) and  $\alpha = \alpha'$  (parallel spins),

$$E_{\uparrow\downarrow}^{(2)} + E_{\downarrow\downarrow}^{(2)}. \quad (7.84)$$

For anti-parallel spins, we obtain

$$\frac{1}{\Omega} E_{\uparrow\downarrow}^{(2)} = \frac{1}{\Omega} \left( \frac{4\pi\alpha}{\Omega} \right)^2 \frac{2m}{\hbar^2} 4 \sum_{\vec{i}\vec{j}\vec{a}\vec{b}} (|\vec{i}|^2 - |\vec{a}|^2 + |\vec{j}|^2 - |\vec{b}|^2)^{-1} \left[ (\mu^2 + |\vec{i} - \vec{a}|^2)^{-1} \right]^2 \delta_{\vec{i} + \vec{j}, \vec{a} + \vec{b}} \quad (7.85)$$

the factor 4 comes from identification of several identical contributions. We are interested in showing that this energy diverges. The proof for the parallel spin case is similar, and will *not* cancel the divergence in  $E_{\uparrow\downarrow}^{(2)}$ . (The eager student can study the parallel spin case in Raimis.)

To get rid of the Kronecker delta, which expresses momentum conservation, we introduce the momentum vector  $\vec{q} \equiv \vec{a} - \vec{i}$ . We then obtain from momentum conservation

$$\vec{a} = \vec{i} + \vec{q}, \quad \vec{b} = \vec{j} - \vec{q}. \quad (7.86)$$

The summation over  $\vec{a}$  and  $\vec{b}$  is replaced by a single summation over  $\vec{q}$ . We introduce integrals, obtaining

$$\begin{aligned} \frac{1}{\Omega} E_{\uparrow\downarrow}^{(2)} &= C \int_{|\vec{i}| < k_F} d\vec{i} \int_{|\vec{j}| < k_F} d\vec{j} \int_{\vec{q}} d\vec{q} \theta(|\vec{j} - \vec{q}| - k_F) \theta(|\vec{i} + \vec{q}| - k_F) \\ &\quad \times (|\vec{i}|^2 + |\vec{j}|^2 - |\vec{i} + \vec{q}|^2 - |\vec{j} - \vec{q}|^2)^{-1} (\mu^2 + |\vec{q}|^2)^{-2}. \end{aligned} \quad (7.87)$$

where  $C$  is a constant independent of  $\Omega$ . The theta function is defined by  $\theta(x) = 0$  if  $x < 0$ , and  $\theta(x) = 1$  if  $x > 0$ . Notice that all powers of  $\Omega$  have been canceled, leaving an integral as a function of  $\mu$  only.

Let us study the integrand, and let us assume that  $\mu$  is very small. The integrand then gets its main contribution from small  $q = |\vec{q}|$ . To see this, note that

$$|\vec{i}|^2 + |\vec{j}|^2 - |\vec{i} + \vec{q}|^2 - |\vec{j} - \vec{q}|^2 = i^2 + j^2 - a^2 - b^2 < 0 \quad (7.88)$$

becomes closest to 0 when  $\vec{q}$  is small. Similarly  $(q^2 + \mu^2)$  is the smallest when  $q$  is small. Thus the integrand has its largest values for  $q$  small.

When  $q$  is small,  $i \lesssim k_F$ : otherwise it is not possible that  $|\vec{a}| = |\vec{i} + \vec{q}| > k_F$ .

$$|\vec{i} + \vec{q}|^2 = i^2 + q^2 + 2i q x > k_F^2, \quad (7.89)$$

where

$$x \equiv \frac{\vec{i} \cdot \vec{q}}{iq} \equiv \cos(\theta_i), \quad (7.90)$$

i.e., we have introduced spherical coordinates for  $\vec{i}$ , aligning the  $z$ -axis along  $\vec{q}$ . For future reference we also define

$$y \equiv -\frac{\vec{j} \cdot \vec{q}}{jq} \equiv -\cos(\theta_j). \quad (7.91)$$

Here,  $\theta_i$  ( $\theta_j$ ) is the angle between  $\vec{q}$  and  $\vec{i}$  ( $\vec{j}$ ). Now, for small  $q$ , we can do a first-order consideration, neglect terms of order  $q^2$  and higher, and a geometrical consideration shows that  $i \approx k_F(1 - cq) + O(q^2)$  for some small number  $c > 0$ . Equation (7.89) becomes, to first order in  $q$ ,

$$k_F^2(1 - 2cq) + 2k_Fqx > k_F^2. \quad (7.92)$$

Rearranging, we obtain

$$qx > (k_F - i), \quad (7.93)$$

and thus the function  $\theta(|\vec{i} + \vec{q}| - k_F)$  is, for small  $q$ , equivalent to the integration limits

$$i \in [k_F - q \cos(\theta_i), k_F], \quad \text{and} \quad \cos(\theta_i) \geq 0. \quad (7.94)$$

A similar analysis for  $|\vec{j} + \vec{q}| > k_F$  gives the integration limits

$$j \in [k_F + q \cos(\theta_j), k_F], \quad \text{and} \quad \cos(\theta_j) \leq 0. \quad (7.95)$$

We write down the integral contribution from  $q \leq \epsilon$  ( $\ll \mu \ll k_F$ ),

$$F(\epsilon, \mu) := \int_{q < \epsilon} \frac{1}{(\mu^2 + q^2)^2} \left[ \int_{A(\vec{q})} d\vec{i} \int_{B(\vec{q})} d\vec{j} \frac{1}{2q(yj + xi)} \right] d\vec{q} \quad (7.96)$$

Here,  $A(\vec{q})$  and  $B(\vec{q})$  denote the integration limits in Eqs. (7.94) and (7.95).

We now introduce spherical coordinates in  $\vec{i}$ -space, letting the  $z$ -axis point in the direction of  $\vec{q}$ . Thus, the elevation angle  $\theta = \theta_i$ , while the azimuthal angle is  $\varphi \in [0, 2\pi]$ . The integration over  $A(\vec{q})$  can be written

$$\int_{A(\vec{q})} d^3\vec{i} = \int_0^{2\pi} d\varphi \int_0^{\pi/2} \sin(\theta) d\theta \int_{k_F - q \cos(\theta)}^{k_F} i^2 di, \quad (7.97)$$

where  $\cos(\theta) = \cos(\theta_i) \geq 0$  is enforced by integrating over  $[0, \pi/2]$ . Reintroduce  $x = \cos(\theta)$ , to obtain

$$\int_{A(\vec{q})} d^3\vec{i} = 2\pi \int_0^1 dx \int_{k_F - qx}^{k_F} i^2 di, \quad (7.98)$$

Similarly, using  $y = -\sin(\theta_j)$  with  $\theta_j \in [\pi/2, \pi]$  to enforce the limits,

$$\int_{B(\vec{q})} d^3\vec{j} = 2\pi \int_0^1 dy \int_{k_F - qy}^{k_F} j^2 dj. \quad (7.99)$$

We now note that in the integration region,  $j^2 \approx k_F^2$ ,  $i^2 \approx k_F^2$ , and thus

$$\frac{i^2 j^2}{yi + xj} \approx \frac{k_F^3}{y + x}, \quad (7.100)$$

the error being small and causing no problems. Thus, to within an error of order  $\epsilon$ ,

$$\begin{aligned} F(\epsilon, \mu) &= 4\pi^2 \int_{|\vec{q}| < \epsilon} \frac{d\vec{q}}{(\mu^2 + q^2)^2 2q} \int_0^1 dx \int_0^1 dy \int_{k_F - qx}^{k_F} di \int_{k_F - qy}^{k_F} dj \frac{k_F^3}{x + y} \\ &= k_F^3 4\pi^2 \int_{|\vec{q}| < \epsilon} \frac{d\vec{q} q^2}{(\mu^2 + q^2)^2 2q} \int_0^1 dx \int_0^1 dy \frac{xy}{x + y}. \end{aligned} \quad (7.101)$$

The integral over  $x$  and  $y$  yields a constant independent of  $\vec{q}$ , thus, noting that the remaining  $\vec{q}$ -integrand does only depend on the magnitude  $q$ ,

$$F(\epsilon, \mu) \propto \int_0^\epsilon \frac{q^3}{(\mu^2 + q^2)^2} dq. \quad (7.102)$$

Consider indefinite integral

$$\int \frac{q^3}{(\mu^2 + q^2)^2} dq = \frac{1}{2} \left[ \frac{\mu^2}{\mu^2 + q^2} + \log(\mu^2 + q^2) \right]. \quad (7.103)$$

For  $\mu > 0$  we see that  $F(\epsilon, \mu)$  is finite, but that *the limit  $\mu \rightarrow 0$  is infinite*.

In total, we see that there is an infinite contribution to  $E^{(2)}$  in the physical limit  $\mu \rightarrow 0$ .

It can be shown that *all*  $E^{(n)}/\Omega$  for  $n \geq 2$  diverge in a similar manner in the thermodynamic limit. Thus, RSPT fails badly for the electron gas. This is not to say that the ground-state energy is not well-defined! One can prove that the energy per particle must be finite in the thermodynamic limit. What fails here are the conditions for RSPT to converge. A necessary condition for convergence is that when we introduce a coupling constant  $\lambda$ , the ground-state energy of  $\hat{H}(\lambda) = \hat{T} + \lambda \hat{W}_0$  is analytic at  $\lambda = 0$ . The infinite perturbation series terms contradicts this assumption.

## Chapter 8

# Common basis sets

In large *ab initio* calculations for realistic systems, one needs a single-particle basis set in which we develop the wavefunction. In this brief chapter, we give a brief overview of the various common choices in quantum chemistry, in solid-state physics, quantum dot studies, and in nuclear physics.

Using  $L$  single-particle functions, the Hilbert space scales as  $\binom{N}{L}$ . Thus, our basis needs to

1. Yield a good approximation to the exact wavefunction, i.e., capture “the physics”,
2. Allow efficient calculation of two-body (or higher!) matrix elements.

These criteria are not always compatible.



## 8.1 Harmonic oscillator basis functions

In several models, the  $N$ -body Hamiltonian can be written

$$\hat{H} \approx \sum_{i=1}^N \hat{h}_{\text{HO}}(i) + \hat{W}, \quad (8.1)$$

where  $W$  is a residual part, and where  $\hat{h}_{\text{HO}}$  is the *harmonic oscillator* (HO),

$$\hat{h}_{\text{HO}} = -\frac{\hbar^2}{2m} \nabla^2 + \frac{1}{2} m \omega^2 r^2 = \sum_{j=1}^d h(r_j). \quad (8.2)$$

Here  $d$  is the spatial dimension of the single-particle space, where the fermions “live”. Typically,  $d = 1, 2$ , or 3. The HO can be exactly solved, giving a convenient basis of single-particle orbitals for manybody treatments.

Harmonic oscillator functions are useful in quantum dot models and very common in the nuclear manybody problem as well.

Consider first a harmonic oscillator in one space dimension (no spin). This simple problem has the Hamiltonian

$$\hat{h} = -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \frac{1}{2} m \omega^2 x^2. \quad (8.3)$$

The the solutions to the eigenvalue problem  $\hat{h}f_n = e_n f_n$  are well-known, and on the form

$$f_n(x) = \sqrt{\frac{\alpha}{2^n n!}} \pi^{-1/4} H_n(\alpha x) e^{-\alpha^2 x^2 / 2} \quad (8.4)$$

where

$$\alpha = \left( \frac{m\omega}{\hbar} \right)^{1/2}. \quad (8.5)$$

The eigenvalues are

$$e_n = \hbar \omega \left( n + \frac{1}{2} \right). \quad (8.6)$$

The functions  $H_n(x)$  are the Hermite polynomials, which have the compact expression

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}. \quad (8.7)$$

The Wikipedia page has tons of information.

### 8.1.1 $d$ -dimensional HO in Cartesian coordinates

The  $d$ -dimensional HO can be solved with separation of variables, since  $\hat{h}_{\text{HO}} = \hat{h}(r_1) + \cdots \hat{h}(r_d)$ . The eigenfunctions become

$$\varphi_{n_1 \cdots n_d}(\vec{r}) = f_{n_1}(r_1) f_{n_2}(r_2) \cdots f_{n_d}(r_d) \quad (8.8)$$

with eigenvalues

$$e_{n_1 \cdots n_d} = \hbar \omega \left( n_1 + n_2 + \cdots + n_d + \frac{d}{2} \right). \quad (8.9)$$

The eigenfunctions are orthonormal,

$$\langle \varphi_{\vec{n}} | \varphi_{\vec{n}'} \rangle = \delta_{\vec{n}, \vec{n}'}. \quad (8.10)$$

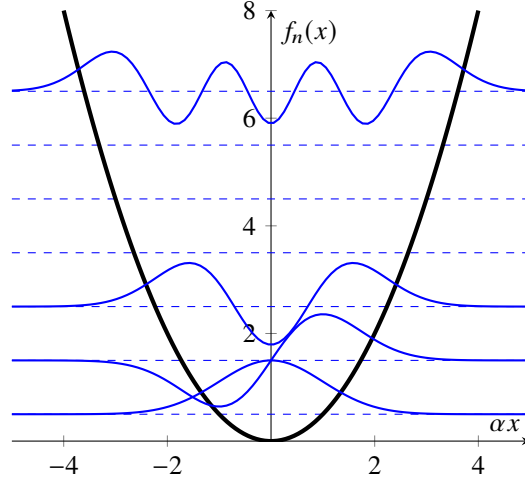


Figure 8.1: The first few Hermite functions  $f_n(x)$ ,  $n = 0, 1, 2$  and  $n = 6$ . They are shifted vertically with their energy eigenvalue. The first eigenvalues are also shown as dashed lines.

For  $d = 2$  we obtain the eigenfunctions

$$\varphi_{n_1, n_2}(\vec{r}) = \alpha \left[ \pi 2^{n_1+n_2} n_1! n_2! \right]^{-1/2} H_{n_1}(\alpha x_1) H_{n_2}(\alpha x_2) e^{-\alpha^2 r^2/2}, \quad r^2 = r_1^2 + r_2^2. \quad (8.11)$$

For  $d = 3$ , we obtain

$$\varphi_{n_1, n_2, n_3}(\vec{r}) = \alpha^{3/2} \pi^{-3/4} \left[ 2^{n_1+n_2+n_3} n_1! n_2! n_3! \right]^{-1/2} H_{n_1}(\alpha x_1) H_{n_2}(\alpha x_2) H_{n_3}(\alpha x_3) e^{-\alpha^2 r^2/2}, \quad r^2 = r_1^2 + r_2^2 + r_3^2. \quad (8.12)$$

The energy levels group into *shells* of equal energy. Define the *shell number*  $N(\vec{n}) = \sum_j n_j$ , such that

$$e_{\vec{n}} = \hbar \omega (N(\vec{n}) + d/2). \quad (8.13)$$

The degeneracy  $g(N, d)$  of the energy  $e = \hbar \omega (N + d/2)$  depends on the dimension  $d$ . Let us look at some examples. Suppose that  $N = 2$ , and  $d = 2$ . Then, we have the possibilities

$$(n_1, n_2) \in \{(0, 2), (2, 0), (1, 1)\}, \quad (8.14)$$

giving a degeneracy of  $g(2, 2) = 3$ . For  $d = 3$  we obtain

$$(n_1, n_2, n_3) \in \{(0, 0, 2), (0, 2, 0), (2, 0, 0), (0, 1, 1), (1, 0, 1), (1, 1, 0)\}, \quad (8.15)$$

giving  $g(2, 3) = 6$ .

In general,  $g(N, 1) = 1$ , and one can show (exercise!) that  $g(N, 2) = N + 1$ , while for  $g(N, 3) = \frac{1}{2}(N+1)(N+2)$ .

In  $d = 2, 3$  one can also find the eigenfunctions of the HO using *polar coordinates*. The main observation is that  $\hat{h}_{\text{HO}}$  is *rotationally invariant*, meaning that it commutes with the generators for the group of space rotations: the angular momentum operators.

**Exercise 8.1.** For the Harmonic oscillator, compute the degeneracy of the eigenvalue  $\hbar \omega (N + d/2)$  for  $d = 1$ ,  $d = 2$ , and  $d = 3$ .  $\triangle$

**Exercise 8.2.** Using your method of choice, plot the Cartesian coordinate eigenfunctions for  $d = 2$  for  $N \leq 3$ . (This constitutes 10 plots.) Set  $\alpha = 1$ .  $\triangle$

### 8.1.2 Polar coordinate HO eigenfunctions, $d = 2$

The case  $d = 2$ : Polar coordinates are defined by

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} r \cos \phi \\ r \sin \phi \end{pmatrix}. \quad (8.16)$$

The group of rotations is generated by the angular momentum operator

$$\hat{L}_z = -i\hbar \frac{\partial}{\partial \phi}. \quad (8.17)$$

The Laplace operator can be written

$$\nabla^2 = \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2}. \quad (8.18)$$

The HO Hamiltonian for  $d = 2$  becomes

$$\hat{h}_{\text{HO}} = -\frac{\hbar^2}{2m} \left[ \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{1}{r^2} \frac{\partial^2}{\partial \phi^2} \right] + \frac{1}{2} m \omega^2 r^2. \quad (8.19)$$

Attempting an eigenfunction of  $\hat{h}_{\text{HO}}$  on the form

$$\varphi(r, \phi) = R(r)u(\phi), \quad (8.20)$$

we obtain after some simple algebra the solution  $u(\phi) = e^{il_z \phi}$  with  $l_z \in \mathbb{Z}$ , and the *radial equation*

$$\left[ -\frac{\hbar^2}{2m} \frac{1}{r} \frac{\partial}{\partial r} r \frac{\partial}{\partial r} + \frac{\hbar^2}{2m} \frac{l_z^2}{r^2} + \frac{1}{2} m \omega^2 r^2 \right] R(r) = e R(r). \quad (8.21)$$

We note that  $e^{il_z \phi}$  are eigenfunctions of  $\hat{L}_z$  with eigenvalue  $\hbar l_z$ .

The radial equation can also be solved, giving  $R_{nl_z}(r)$ ,  $n = 0, 1, \dots$ , and a total normalized eigenfunction

$$\varphi_{nl_z}(r, \phi) = \alpha \left[ \frac{2n!}{2\pi(|l_z| + n)!} \right]^{1/2} (\alpha r)^{|l_z|} e^{-\alpha^2 r^2/2} L_n^{|l_z|}(\alpha^2 r^2) e^{il_z \phi}, \quad (8.22)$$

where  $L_n^k(x)$  are the associated Laguerre polynomials. These are given by

$$L_n^k(x) = \frac{1}{n!} e^x x^{-k} \frac{d^n}{dx^n} (e^{-x} x^{n+k}), \quad (8.23)$$

and are polynomials of degree  $n$ . It should be observed that the space part of  $\varphi_{nl_z}$  is a Gaussian multiplied with a polynomial of degree  $2n + |l_z|$  (in  $r$ ). The energy eigenvalue of  $\varphi_{n,l_z}$  is given by

$$e_{nl_z} = \hbar\omega(2n + |l_z| + 1), \quad (8.24)$$

and we see that each shell is given by  $N = 2n + |l_z|$ . It is a fact that  $L_n^{|l_z|}$  has  $n$  nodes (which are all nonzero). Thus,  $n$  counts the nodes of the radial wavefunction except for  $r = 0$ , which is a node for  $l_z \neq 0$ .

Note well that the quantum numbers  $(n, l_z)$  are *not identical* to the Cartesian coordinate quantum numbers  $(n_1, n_2)$  used previously, even though our notation for the eigenfunctions is the same.

**Exercise 8.3.** Using your method of choice, plot the polar coordinated eigenfunctions for  $d = 2$  for  $N = 2n + |l_z| \leq 3$ . (This constitutes 10 plots.) Set  $\alpha = 1$ .  $\triangle$

### 8.1.3 Polar coordinate HO eigenfunctions, $d = 3$

For  $d = 3$ ,  $[\hat{h}, \hat{L}_z] = [\hat{h}, \hat{L}^2] = [\hat{L}_z, \hat{L}^2] = 0$ , so we can find a common set of eigenvectors for the three operators  $\hat{h}$ ,  $\hat{L}_z$ , and  $\hat{L}^2$ .

Polar coordinates in 3D (aka spherical coordinates) are defined by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} r \sin \theta \cos \phi \\ r \sin \theta \sin \phi \\ r \cos \theta \end{pmatrix}, \quad r \in [0, +\infty), \quad \theta \in [0, \pi], \quad \phi \in [0, 2\pi]. \quad (8.25)$$

The Laplacian in polar coordinates becomes

$$\nabla^2 = \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} - \frac{1}{\hbar^2 r^2} \hat{L}^2, \quad (8.26)$$

where  $\hat{L}^2 = \hat{L}_x^2 + \hat{L}_y^2 + \hat{L}_z^2$  is, in polar coordinates,

$$\hbar^{-2} \hat{L}^2 = -\frac{1}{\sin \theta} \frac{\partial}{\partial \theta} \sin \theta \frac{\partial}{\partial \theta} - \frac{1}{\sin^2 \theta} \frac{\partial^2}{\partial \phi^2}. \quad (8.27)$$

As for the  $d = 2$  case, we attempt an eigenfunction of  $\hat{h}_{\text{HO}}$  on the form  $\varphi(r, \theta, \phi) = R(r)Y(\theta, \phi)$ . Straightforward algebra leads to  $Y$  being an eigenfunction of  $\hat{L}^2$ . The spherical harmonics  $Y_{ll_z}(\theta, \phi)$  form a complete set of eigenfunctions of  $\hat{L}^2$  (and  $\hat{L}_z$ ),

$$\hat{L}^2 Y_{ll_z}(\theta, \phi) = \hbar^2 l(l+1) Y_{ll_z}(\theta, \phi), \quad (8.28)$$

$$\hat{L}_z Y_{ll_z}(\theta, \phi) = \hbar l_z Y_{ll_z}(\theta, \phi). \quad (8.29)$$

The explicit expression is

$$Y_{ll_z}(\theta, \phi) = \left[ \frac{2l+1}{4\pi} \frac{(l-l_z)!}{(l+l_z)!} \right]^{1/2} P_l^{l_z}(\cos \theta) e^{il_z \phi}, \quad |l_z| \leq l \in \mathbb{N}. \quad (8.30)$$

The radial equation becomes

$$\left[ -\frac{\hbar^2}{2m} \frac{1}{r^2} \frac{\partial}{\partial r} r^2 \frac{\partial}{\partial r} + \frac{\hbar^2 l(l+1)}{2mr^2} + \frac{1}{2} m \omega^2 r^2 \right] R(r) = e R(r). \quad (8.31)$$

Note that the radial equation depends on  $l$ , but not on  $l_z$ . The radial equation can be solved, giving (see Moshinsky's book [9]), solutions  $R_{nl}(r)$  for  $n = 0, 1, 2, \dots$ ,

$$R_{nl}(r) = \alpha^{3/2} \left[ \frac{2(n!)}{\Gamma(n+l+3/2)} \right]^{1/2} (\alpha r)^l L_n^{l+1/2}(\alpha^2 r^2) e^{-\alpha r^2/2}. \quad (8.32)$$

(Remark: in some texts, a different convention for  $n$  is used. Note carefully that  $l$  is not restricted with respect to  $n$ .)

The HO energy is

$$e_{nl} = \hbar \omega (2n + l + 3/2), \quad (8.33)$$

which is *independent of  $l_z$*  since the radial equation was independent of  $l_z$ .

**Exercise 8.4.** Using your method of choice, plot the polar coordinated eigenfunctions for  $d = 3$  for  $N = 2n + l \leq 3$ . Set  $\alpha = 1$ . Use only  $l_z \geq 0$ .  $\triangle$

## 8.2 Plane-wave basis set

Already covered

## 8.3 The nuclear manybody problem in a non-relativistic approximation

The atomic nucleus consists of neutrons  $n$  and protons  $p$ . These are compound particles, consisting of three quarks each. The neutron consists of 1 up quarks of charge  $+(2/3)e$ , and 2 down quarks of charge  $-(1/3)e$ . The proton consists of 2 up quarks and 1 down quark. Thus, in total the neutron has no charge, while the proton has charge  $+e$ .

Experimental evidence demonstrates that  $p$  and  $n$  behave almost identically in the nucleus, with almost equal mass, spin  $+\hbar/2$ , and that they do not decompose into their constituent quarks at low energy. Their interactions in a nucleus is also almost identical, except that two protons repel each other via the Coulomb force. Therefore, Werner Heisenberg postulated a non-relativistic description where  $p$  and  $n$  are different states of *one kind of particle*, a *nucleon*, with an additional spin-1/2 degree of freedom called *isospin*: a nucleon with isospin  $+1/2$  is a proton, and a nucleon with isospin  $-1/2$  is a neutron. (It was Eugene Wigner who coined the term “isospin” in 1937.)

Thus, the Hamiltonian of an  $A$ -particle nucleus is

$$\hat{H} = \hat{T} + \hat{U} = \sum_{i=1}^N \hat{t}(i) + \frac{1}{2} \sum_{i \neq j}^A \hat{u}(i, j), \quad (8.34)$$

where  $\hat{u}(i, j)$  is the interaction potential between nucleons  $i$  and  $j$ . Note well, that the interaction potential depends on the isospin of nucleons  $i$  and  $j$ .

The interaction potential is *not known a priori*, in contrast to electronic systems. One needs to fit semi-empirical models to experimental data, or, as is the current trend, derive potential approximations from QCD.

Suppose we are given a spatial orbital basis  $\varphi_p(\vec{r})$ , so, the HO function with  $p = (n, l, l_z)$ . Since we have both spin and isospin in our single-particle space, we obtain single-particle functions on the form

$$\phi_{p,\alpha,t}(\vec{r}, \sigma, \tau) = \varphi_p \chi_\alpha(\sigma) \chi_t(\tau), \quad (8.35)$$

where  $\chi_{\pm 1/2}$  are the orthonormal spin-1/2 basis functions. It can be considered standard to use the HO basis functions for nonrelativistic treatments of the nucleus.

A Slater determinant with  $N$  neutrons and  $Z$  protons, in total  $A = N + Z$  nucleons, can then be written

$$|(p_1\alpha_1, +1/2) \cdots (p_Z\alpha_Z, +1/2)(p_{Z+1}\alpha_{Z+1}, -1/2)(p_{Z+2}\alpha_{Z+2}, -1/2) \cdots (p_A\alpha_A, -1/2)\rangle \quad (8.36)$$

Suppose we have in total  $L$  values for  $(p\alpha)$ , i.e.,  $L$  spin-orbitals. Since the isospin values for neutrons and protons are different, neutrons and protons can occupy spin-orbitals independently of each other, meaning that the dimension of the Hilbert space becomes

$$D = \binom{L}{N} \times \binom{L}{Z}. \quad (8.37)$$

### 8.3.1 The self-bound property of the nucleus, and removal of center-of-mass degree of freedom

The nucleus is *self-bound*. There is no external potential in the one-body part of the Hamiltonian to bind the nucleons in space. The Hamiltonian is *translationally invariant*, i.e., it commutes with the total momentum operator  $\hat{P} = \sum_i \hat{p}(i)$ . The spectrum of  $\hat{H}$  becomes purely continuous, there are no isolated eigenvalues. This complicates matters for most common manybody techniques. It is therefore a common technique to add a weak fictitious harmonic oscillator potential  $\frac{\mu}{\omega} r^2$  term to the Hamiltonian to weakly bind the nucleus, producing a discrete spectrum, and then after the calculations remove the  $\omega$  dependence.

**NB:** To be added, material not lectured: center-of-mass transformation.

## 8.4 Molecular systems and Gaussian basis sets

### 8.4.1 The Born–Oppenheimer molecular Hamiltonian

Classically, a molecule is a collection of nuclei with masses  $M_\alpha$ , charges  $eZ_\alpha$ , and positions  $\vec{R}_\alpha$ ,  $\alpha = 1, 2, \dots, N_{\text{at}}$ , and a collection of  $N$  electrons with charge  $-e$  and positions  $\vec{r}_i$ ,  $i = 1, 2, \dots, N$ . Quantum mechanically, both the nuclei and the electrons obtain spin, and a wavefunction depending on *all* the  $N + N_{\text{at}}$  space-spin coordinates. In all but the simplest cases, this is an intractable problem.

The way out is the *Born–Oppenheimer (BO) approximation*: Roughly speaking<sup>1</sup>, the nuclei are so heavy compared to the electrons that their movement occurs on a time-scale much larger than the motion of the electrons. In the BO approximation we therefore treat the nuclei as classical particles, setting up an external classical electrostatic potential  $v(\vec{r})$  felt by an electron,

$$v(\vec{r}) = \sum_{\alpha=1}^{N_{\text{at}}} \frac{-e^2 Z_\alpha}{|\vec{r} - \vec{R}_\alpha|}. \quad (8.38)$$

The molecular Hamiltonian is therefore an  $N$ -electron Hamiltonian with a *parametric dependence* on the nuclear geometry,

$$\hat{H} = \hat{H}(\vec{R}_1, \vec{R}_2, \dots, \vec{R}_{N_{\text{at}}}). \quad (8.39)$$

The total BO Hamiltonian becomes

$$\hat{H} = \hat{H}_{\text{e-e}} + \hat{H}_{\text{n-e}} + \hat{H}_{\text{n-n}} \quad (8.40)$$

where

$$\hat{H}_{\text{e-e}} = \hat{T} + \hat{W} \quad (8.41)$$

describes the kinetic energy and Coulomb repulsion among the electrons. Furthermore,

$$\hat{H}_{\text{n-e}} = \hat{V} = \sum_{i=1}^N v(\vec{r}_i) = \sum_{i\alpha} \frac{-e^2 Z_\alpha}{|\vec{r}_i - \vec{R}_\alpha|} \quad (8.42)$$

is the interactions between the electrons and the nuclei. Finally,

$$\hat{H}_{\text{n-n}} = \frac{1}{2} \sum_{\alpha \neq \beta=1}^{N_{\text{at}}} \frac{e^2 Z_\alpha Z_\beta}{|\vec{R}_\alpha - \vec{R}_\beta|} \quad (8.43)$$

is a constant term depending on the nuclear geometry. Let us recall the expressions for  $\hat{T}$  and  $\hat{W}$ ,

$$\hat{T} = \sum_{i=1}^N \left( -\frac{\hbar^2}{2m} \nabla_i^2 \right) \quad (8.44)$$

$$\hat{W} = \frac{1}{2} \sum_{i \neq j=1}^N \frac{e^2}{|\vec{r}_i - \vec{r}_j|}. \quad (8.45)$$

The eigenvalues  $E_k$  and eigenfunctions  $|\Psi_k\rangle$  of  $\hat{H}$  obtain a parametric dependence on the nuclear geometry. Of particular usefulness in chemistry is the *potential energy surface*: the ground-state energy  $E_0(\vec{R}_1, \dots, \vec{R}_{N_{\text{at}}})$  as a function of the nuclear coordinates.

<sup>1</sup>The BO approximation has some subtleties, but these are beyond the scope of this course.

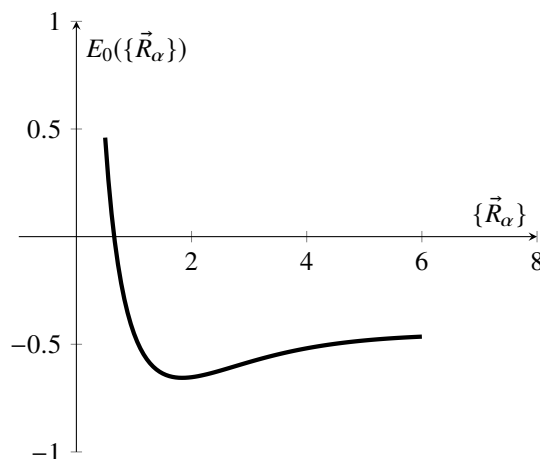


Figure 8.2: Schematic illustration of a potential energy surface.

The *equilibrium geometry* is the configuration of the nuclei that minimizes  $E_0$ . This usually corresponds to the configuration observed in nature.

We will not have more to say on the topic. The interested student should consult for example the book by Szabo and Ostlund [3] – a great read. The book [6] is the definite guide to modern electronic-structure theory.

### 8.4.2 Hartree–Fock and Post-Hartree–Fock methods

It is a fact that Hartree–Fock in most cases works really well for molecular systems. Most of the contributions to the eigenenergies are captured by this approximation. Therefore, it is almost universally accepted in the quantum chemistry community to *first* do a HF calculation (RHF or UHF, but rarely general HF). One *then* introduces a more advanced description using the HF orbitals thus obtained. This places (Møller–Plesset) perturbation theory, CI and CC methods in the category *post-Hartree–Fock methods*.

Since HF is a mean-field model where the wavefunction is on Slater determinant form, it is referred to as “uncorrelated”. One defines *the correlation energy* as the difference

$$\Delta E = E_0 - E_{\text{HF}}. \quad (8.46)$$

The correlation energy  $\Delta E$  is usually, but not always small.

The usual strategy is as follows:

1. Choose a set of *atomic orbitals*  $\chi_p(\vec{r})$ . These may or may not be orthonormalized. Usually, they are *not*, for practical reasons.
2. Evaluate matrix elements of the one-body Hamiltonian  $\hat{H}_0 = \hat{T} + \hat{V}$  and the two-body Hamiltonian  $\hat{W}$ , i.e., obtain  $(\chi_p|\hat{h}|\chi_q)$  and  $(\chi_p\chi_q|\hat{w}|\chi_r\chi_s)$ . This is usually done by library functions that are highly complicated in their own.
3. Perform either a RHF or a UHF (but rarely a general HF) calculation to obtain *molecular orbitals*  $\varphi_p$ . Thus, molecular orbitals (MOs) are usually a synonym for HF orbitals. See the section on RHF in a given basis. The MOs are thus given as linear combinations of AOs,

$$\varphi_p = \sum_q \chi_q U_{qp}, \quad (8.47)$$



and the matrix elements  $(\varphi_p \varphi_q | \hat{W} | \varphi_r \varphi_s)$  are thus given as linear combinations of the matrix elements  $(\chi_p \chi_q | \hat{W} | \chi_r \chi_s)$ .

4. Use the molecular orbitals  $\varphi_p$  in a many-body treatment such as MPPT, CI, or coupled-cluster theory.

### 8.4.3 From hydrogenic to Gaussian orbitals

How do we choose a single-particle basis (atomic orbitals) for the BO Hamiltonian? Clearly, the basis must depend on the nuclear arrangement: we would like our results to be independent of translations of the whole molecule, as this is a fundamental symmetry in the problem. The intuition behind the BO approximation also indicates that each individual *atom* in the molecule roughly retains its independence as an entity in itself: the electron cloud of a molecule is a perturbation of the electron cloud obtained by treating each atom by itself, eliminating inter-atom interactions.

We therefore consider first an individual atom for guidance, located for convenience at  $\vec{R} = 0$ , with nuclear charge  $eZ$ . We assume that the atom has  $N$  electrons, and we first consider the non-interacting problem. We thus need to solve for the eigenvalues and eigenstates of a hydrogen-like atom with a single electron. The Hamiltonian reads

$$\hat{h} = -\frac{\hbar^2}{2m} \nabla^2 - \frac{Ze^2}{r}. \quad (8.48)$$

The diagonalization of this problem is textbook material, see for instance [6] or [3]. One finds a sequence of eigenvalues

$$e_n = -\frac{m}{2\hbar^2} \frac{(Ze^2)^2}{n^2}, \quad n = 1, 2, \dots, \quad (8.49)$$

degenerate in the angular momentum quantum numbers  $l \leq n$  and  $l_z$ ,  $|l_z| \leq l$ . The eigenfunctions are given by

$$\psi_{nlz}(\vec{r}) = R_{nl}(r) Y_{ll_z}(\theta, \phi), \quad R_{nl}(r) = \sqrt{\left(\frac{2}{na_0}\right)^3 \frac{(n-l-1)!}{2n(n+l)!}} e^{-\rho/2} \rho^l L_{n-l-1}^{2l+1}(\rho) \quad (8.50)$$

with

$$\rho = \frac{2Ze^2}{n} \frac{m}{\hbar^2} r, \quad a_0 = \frac{\hbar^2}{mZe^2} \quad (8.51)$$

Intuitively, the functions  $\psi_{nlz}$  should be a good single-particle basis for the interacting  $N$ -electron atom, obtained by throwing  $N - 1$  more electrons into this atom and turning on interactions. However, this basis set has major deficiencies:

- They are *incomplete* (and thus not an actual  $L^2(X)$ -basis!), as the hydrogen atom also has a continuous spectrum for energies  $e > 0$ . Thus, we cannot expect convergence to the exact ground-state energy of the  $N$ -electron atom as we include more and more  $\psi_{nlz}$ .
- Computing the matrix elements of  $\hat{W}$  becomes complicated.
- The functions become very *diffuse* with higher  $n$ , allowing few details to be resolved around the nucleus for moderate basis sizes.

On the other hand, the basis set displays other very useful features in its asymptotic behavior:

- A nuclear *cusp* at the origin, stemming from the singular nature of the Coulomb potential. This cusp is always present in an atom, and gives a large contribution to the total electronic energy.

- Exponential fall-off of the radial part. This is responsible for *physics* of the  $N$ -electron atoms and molecules, such as an  $R^{-1}$ -dependence of the inter-atomic forces in a molecule, where  $R$  is the distance between two atoms.

A partial remedy to the problems is the use of *Laguerre radial functions*, see [6]. These have nuclear cusps and exponential fall-off, while forming a complete set. These functions do not solve the problems of the complicated  $\hat{W}$  matrix elements, however. We will not study these functions in detail here.

#### 8.4.4 Gaussian basis sets

Selecting a single-particle basis for molecular systems is an art, due to the conflicting constraints of efficiency, compactness and accuracy. Moreover, different manybody methods put different requirements on the basis set. There are probably *hundreds* of different basis sets, with acronyms like “STO- $k$ G”, “cc-PVXZ”, etc. They are all tailored to have specific behavior, and to be useful under different conditions. They all have one thing in common, however: they are linear combinations of *Gaussians*.

Thus, the almost universally used approach in quantum chemistry today is a *pragmatic one*: One uses *Gaussian functions* to approximate single-particle basis functions. A Gaussian is a function on the form

$$g_{ijk}(\vec{r}; \zeta) = \mathcal{N} x^i x^j x^k e^{-\zeta r^2}, \quad (8.52)$$

where the exponent  $\zeta > 0$  is a parameter, and where  $\mathcal{N}$  is a normalization constant. These are closely related to the harmonic oscillator eigenfunctions. In fact, the HO eigenfunctions are finite linear combinations of such  $g_{ijk}$ , since  $H_n(x)$  is a polynomial. Conversely, the Gaussians can be expanded in a finite number of HO functions.

The Gaussian  $g_{ijk}(\cdot; \zeta)$  is referred to as a *Cartesian Gaussian* since it is a tensor product of one-dimensional Gaussians  $g_i(x; \zeta) = \mathcal{N} x^i e^{-\zeta x^2}$ .

A more compact description is obtained using *spherical Gaussians* on the form

$$g_{nll_z}^{\text{sph}}(\vec{r}; \zeta) = \mathcal{N} r^l e^{-\zeta r^2} Y_{ll_z}(\theta, \phi). \quad (8.53)$$

These give a more compact description since they are eigenfunctions of  $\hat{L}^2$  and  $\hat{L}_z$ , unlike the Cartesian counterparts. But they are equivalent: the Cartesian and spherical Gaussians can be expanded in terms of each other, using finite number of coefficients.

A general basis of atomic orbitals is then on the form:

$$\chi_p = \sum_{\mu} D_{p\mu} g_{\mu}(\vec{r} - \vec{R}_p; \zeta_{p\mu}), \quad (8.54)$$

$g_{\mu}$  is either the spherical or Cartesian Gaussian functions, and where  $\mu = (ijk)$  or  $\mu = (nll_z)$ . Each  $\chi_p$  needs to be located on some atom  $\vec{R}_p$ . The vector  $\vec{R}_p$  therefore shifts the Gaussian accordingly. The exponents depend on both  $p$  and  $\mu$ , giving maximum flexibility in the description. The matrix  $D$  is typically sparse.

#### 8.4.5 Gaussians are useful because they give fast integration

The reason why Gaussians are almost universally accepted is the fact that we can integrate the Coulomb interaction matrix and the nuclear attraction matrix elements *efficiently*. A large part of the reason is that the product of two Gaussians is easily expressible in terms of other Gaussians – even if they are located at different atoms. Moreover, the actual integration over  $\vec{r}_1$  and  $\vec{r}_2$  in  $(\chi_{\mu} \chi_{\nu} | \hat{w} | \chi_{\mu'} \chi_{\nu'})$  can be carried out semi-analytically in a highly efficient manner. We will not go into details, see [6] for a detailed account of molecular integrals.

## Chapter 9

# Coupled-cluster theory (CC)

Recommended reading: Crawford and Schaefer [10] is a very nice and pedagogical text. Shavitt and Bartlett [11] is also recommended. We mostly follow Crawford and Schaefer here.

This part of the lecture notes is so far rudimentary. Luckily, the Crawford and Schaefer text is *excellent* and our lectures follow this article closely, with only minor changes.

## 9.1 Motivation and introduction: Cluster functions and the exponential ansatz

Consider a Slater determinant ansatz to the  $N$ -fermion wavefunction,

$$|\Phi\rangle = |\phi_1 \phi_2 \cdots \phi_N\rangle, \quad (9.1)$$

where we for simplicity fill the  $N$  first single-particle functions. The fermions in this wavefunction are *uncorrelated*, except for the Pauli principle. It is the simplest manybody ansatz we can make.

We are going to motivate an *exponential ansatz* for the exact wavefunction,

$$|\Psi\rangle = e^{\hat{T}} |\Phi\rangle, \quad (9.2)$$

where  $\hat{T}$  is an excitation operator called a *cluster operator*. The exponential ansatz is *nonlinear*, and will lead to the very popular *coupled-cluster method* (CC).

Assume that  $|\Phi\rangle$  is a *reasonable* ansatz for the exact wavefunction  $|\Psi\rangle$ , i.e., that at least  $\langle\Phi|\Psi\rangle \neq 0$ . By scaling  $|\Psi\rangle$  by a number, we can write

$$|\Psi\rangle = |\Phi\rangle + |\Delta\Psi\rangle, \quad (9.3)$$

How can we improve on  $|\Phi\rangle$  in a systematic manner towards  $|\Psi\rangle$ ? The Slater determinant is an antisymmetrized tensor product,

$$\Phi(1, 2, \dots, N) = \sqrt{N!} \mathcal{A} \phi_1(1) \phi_2(2) \cdots \phi_N(N). \quad (9.4)$$

The antisymmetrizer  $\mathcal{A}$  was defined in Exercise 2.5 on page 17, and is the orthogonal projection from  $L^2(X^N)$  to  $L^2(X^N)_{\text{AS}}$ . The action of  $\mathcal{A}$  on any unsymmetrized wavefunction will give a properly antisymmetric function.

Intuitively, if we add to the *product*  $\phi_1(1)\phi_2(2)$  a general function  $g_{12}(1, 2)$ , we would obtain a wavefunction where “2 of the fermions are correlated”, i.e., described with a general wavefunction, while the rest are still independent,

$$\begin{aligned} \Psi_{\text{better}}(1, 2, \dots, N) &= \sqrt{N!} \mathcal{A} [\phi_1(1)\phi_2(2) + g_{12}(1, 2)] \phi_3(3) \cdots \phi_N(N) \\ &\equiv \Phi(1, 2, \dots, N) + \langle 12 \cdots N | g_{12} \phi_3 \cdots \phi_N \rangle \end{aligned} \quad (9.5)$$

The antisymmetrization operator  $\mathcal{A}$  ensures that the final wavefunction is fully antisymmetrized. The latter equation defines  $|g_{12}\phi_3 \cdots \phi_N\rangle$  via the antisymmetrization operation on a product. Thus, in ket notation,

$$|\Psi_{\text{better}}\rangle = |\Phi\rangle + |g_{12}\phi_3 \cdots \phi_N\rangle. \quad (9.6)$$

The function  $g_{12}(x, y)$  is called a *cluster function*, since when applied to  $|\Phi\rangle$  in the above manner it describes the wavefunction of a system where all fermions are independent/uncorrelated (think far away from each other), except for a single *cluster* of particles consisting of 2 fermions that are described in a general manner (close to each other).<sup>1</sup>

Suppose we instead introduce a correction on the occupied SPFs  $\phi_i$  and  $\phi_j$ ,  $i < j$ .

$$\begin{aligned} \Psi'_{\text{better}}(1, 2, \dots, N) &= \sqrt{N!} \mathcal{A} [\phi_i(i)\phi_j(j) + g_{ij}(i, j)] \phi_1(1) \cdots \cancel{\phi_i(i)} \cdots \cancel{\phi_j(j)} \cdots \phi_N(N) \\ &= \Phi(1, 2, \dots, N) + (-1)^{i-j+1} \langle 12 \cdots N | g_{ij} \phi_1 \cdots \cancel{\phi_i} \cdots \cancel{\phi_j} \cdots \phi_N \rangle \end{aligned} \quad (9.7)$$

<sup>1</sup>The  $N$  fermions in the system are identical. Hence, it is not meaningful to say “one fermion is here”, or “ $n$  fermions are there”, since, in a sense, *all*  $N$  fermions are involved in such statements. Instead, one speaks of *clusters* of  $n$  fermions. This term is then subtly different from a subset of the fermions.

Here, we used antisymmetry of Slater determinants to find an expression for the correlated part, since  $i < j$  are not necessarily next to each other. (However, note that  $|\phi_1 \cdots g_{i,i+1} \cdots \phi_N\rangle$  is well-defined.)

An even better approach would be to introduce cluster functions  $g_{ij}$  for *all pairs of occupied SPFs*, and in *all possible ways correlate pairs of SPFs*. For example, for  $N = 4$  for simplicity,

$$\begin{aligned} |\Psi_{\text{CCD}}\rangle = & |\phi_1\phi_2\phi_3\phi_4\rangle + |g_{12}\phi_3\phi_4\rangle + |\phi_1g_{23}\phi_4\rangle + |\phi_1\phi_2g_{34}\rangle - |\phi_1\phi_3g_{24}\rangle - |g_{13}\phi_2\phi_4\rangle + |g_{14}\phi_2\phi_3\rangle \\ & + |g_{12}g_{34}\rangle - |g_{13}g_{24}\rangle + |g_{14}g_{23}\rangle \end{aligned} \quad (9.8)$$

This is the coupled-cluster doubles (CCD) wavefunction, for  $N = 4$ . The terms with two cluster functions are defined in a similar way as the terms with only one cluster function.

The function  $g_{ij}(x, y)$  of two one-particle coordinates,  $x, y \in X$ , can be expanded in the SPFs,

$$g_{ij}(x, y) = \sum_{p < q} t_{ij}^{pq} \phi_p(x) \phi_q(y). \quad (9.9)$$

We sum only over  $p < q$ , because we will see that in the end the other coefficients are not independent, by antisymmetry properties of the wavefunction. Inserting this expansion leads to, for  $ij = 12$ ,

$$\langle 12 \cdots N | g_{12} \phi_3 \cdots \phi_N \rangle = \sum_{p < q} t_{12}^{pq} \sqrt{N!} \mathcal{A} \phi_p(1) \phi_q(2) \phi_3(3) \cdots \phi_N(N). \quad (9.10)$$

We now observe that the right-hand side is a linear combination of Slater determinants,

$$\begin{aligned} |g_{12} \phi_3 \cdots \phi_N\rangle &= \sum_{p < q} t_{12}^{pq} |\phi_p \phi_q \phi_3 \cdots \phi_N\rangle \\ &= \sum_{a < b} t_{12}^{ab} c_a^\dagger c_1 c_b^\dagger c_2 |\Phi\rangle \end{aligned} \quad (9.11)$$

In the last equality, we used the fact that only if  $pq = ab$  (virtual SPFs) can we get contributions, due to antisymmetry of Slater determinants. We now note, that including  $a > b$  in the summation does not lead to independent terms, justifying the restriction  $a < b$  in the summation.

Similarly, for the correction term in  $|\Psi'_{\text{better}}\rangle$ ,

$$\begin{aligned} (-1)^{i-j+1} |g_{ij} \phi_1 \cdots \cancel{\phi_i} \cdots \cancel{\phi_j} \cdots \phi_N\rangle &= \sum_{p < q} t_{ij}^{pq} (-1)^{i-j+1} |\phi_p \phi_q \phi_1 \cdots \cancel{\phi_i} \cdots \cancel{\phi_j} \cdots \phi_N\rangle \\ &= \sum_{p < q} t_{ij}^{pq} |\phi_1 \cdots \phi_p \cdots \phi_q \cdots \phi_N\rangle \\ &= \sum_{a < b} t_{ij}^{ab} c_a^\dagger c_i c_b^\dagger c_j |\Phi\rangle. \end{aligned} \quad (9.12)$$

We see that it is useful to define *cluster operators*,

$$\hat{t}_{ij} \equiv \sum_{a < b} t_{ij}^{ab} c_a^\dagger c_i c_b^\dagger c_j, \quad (9.13)$$

and we observe that

$$|\Psi_{\text{better}}\rangle = |\Phi\rangle + \hat{t}_{12} |\Phi\rangle, \quad (9.14)$$

and similarly,

$$|\Psi'_{\text{better}}\rangle = |\Phi\rangle + \hat{t}_{ij} |\Phi\rangle. \quad (9.15)$$

We now notice something curious and important: all operators  $\hat{t}_{ij}$  commute among themselves. Why? They are linear combinations of products of excitation operators  $c_a^\dagger c_i$ , and these commute:

$$[c_a^\dagger c_i, c_{a'}^\dagger c_{i'}] = 0, \quad (9.16)$$

since the creation operators always refer to virtual SPFs and the annihilation operators to occupied SPFs.

**Exercise 9.1.** Prove Eq. (9.16). △

Using this fact, we can then write

$$|\Psi_{\text{CCD}}\rangle = |\Phi\rangle + \sum_{i<j} \hat{t}_{ij} |\Phi\rangle + \frac{1}{2} \sum_{i<j} \hat{t}_{ij} \sum_{i'<j'} \hat{t}_{i'j'} |\Phi\rangle. \quad (9.17)$$

The reader should check that this final equation actually reproduces Eq. (9.8). The factor 1/2 in the last term stems from double-counting of the cluster operators.

**Exercise 9.2.** Prove that Eq. (9.8) becomes Eq. (9.17) when using the definition of  $\hat{t}_{ij}$ . △

We simplify further. If we define the *doubles cluster operator*

$$\hat{T}_2 = \sum_{i<j} \hat{t}_{ij} = \frac{1}{4} \sum_{ij} \sum_{ab} t_{ij}^{ab} c_a^\dagger c_i c_b^\dagger c_j, \quad (9.18)$$

introducing antisymmetry of the *amplitudes*  $t_{ij}^{ab}$ , we have

$$|\Psi_{\text{CCD}}\rangle = (1 + \hat{T}_2 + \frac{1}{2} \hat{T}_2^2) |\Phi\rangle. \quad (9.19)$$

We now observe that in the function  $\hat{T}_2^2 |\Phi\rangle$ , no SPFs with indices  $i \leq N$  are left, since  $N =$ . Thus,  $\hat{T}_2^3 |\Phi\rangle = 0$ , and we have in fact

$$|\Psi_{\text{CCD}}\rangle = e^{\hat{T}_2} |\Phi\rangle. \quad (9.20)$$

The choice  $N = 4$  is not special: for any  $N$ , the wavefunction  $|\Psi_{\text{CCD}}\rangle = e^{\hat{T}_2} |\Phi\rangle$  is identical to the wavefunction where we replace pairs of occupied SPFs by with pair cluster functions  $g_{ij}$  in all possible ways in the reference Slater determinant  $|\Phi\rangle$ .

Furthermore, there is nothing special about *pair* clusters. We may introduce a singles cluster operator

$$\hat{T}_1 = \sum_i \hat{t}_i = \sum_{ia} t_i^a c_a^\dagger c_i, \quad (9.21)$$

corresponding to adding to the various  $\phi_i$  the SPF  $g_i = \sum_a t_i^a \phi_a$ . We may also introduce a triples cluster operator

$$\hat{T}_3 = \sum_{i<j<k} \hat{t}_{ijk} = \frac{1}{3!} \sum_{ijk} \sum_{abc} t_{ijk}^{abc} c_a^\dagger c_i c_b^\dagger c_j c_c^\dagger c_k, \quad (9.22)$$

correlating a cluster of *three* particles, by adding to  $\phi_i \phi_j \phi_k$  a function  $g_{ijk}(x, y, z)$ .

We define a general cluster operator

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \hat{T}_3 + \cdots + \hat{T}_N. \quad (9.23)$$

The expansion stops at  $\hat{T}_N$  because it is impossible to do further corrections!

The wavefunction

$$|\Psi\rangle = e^{\hat{T}} |\Phi\rangle \quad (9.24)$$

is the most general wavefunction obtained from  $|\Phi\rangle$  by correlating all 1 particle clusters, 2 particle clusters, etc, in *all possible ways*. The exponential ansatz is thus called *the cluster expansion*.

In fact, as we will show, *any* wavefunction  $|\Psi\rangle$  with  $\langle\Phi|\Psi\rangle = 1$  can be written as

$$|\Psi\rangle = e^{\hat{T}} |\Phi\rangle . \quad (9.25)$$

Thus, *the cluster expansion represents a systematic way to improve upon the reference wavefunction  $|\Phi\rangle$* . The parameters of this expansion are the cluster amplitudes  $t_i^a$ ,  $t_{ij}^{ab}$ , etc, and they occur in a *nonlinear fashion*.

What is so good about this particular systematic expansion of  $|\Psi\rangle$ ? The answer is *size-consistency*. We will have more to say about this later. However, here is a handwaving argument: Consider the CCD wavefunction for  $N = 4$ , as above. Suppose that  $\phi_1$  and  $\phi_2$  have very small overlap with  $\phi_3$  and  $\phi_4$ . Since  $|\Phi\rangle$  is supposed to be a *reasonable* guess for  $|\Psi\rangle$ , this means that the fermions form 2-fermion clusters that are “far apart”. It is therefore reasonable that  $g_{12}$  and  $g_{34}$  are the only contributing cluster functions to  $|\Psi_{\text{CCD}}\rangle$ : all the other  $g_{ij}$  couple clusters that are very far apart and are approximately zero. We obtain

$$|\Psi_{\text{CCD}}\rangle \approx |\phi_1\phi_2\phi_3\phi_3\rangle + |g_{12}\phi_3\phi_4\rangle + |\phi_1\phi_2g_{34}\rangle + |g_{12}g_{34}\rangle . \quad (9.26)$$

The last term comes from  $\frac{1}{2}\hat{T}_2^2$  – a *quadruples cluster operator*. Compare this with the CI doubles wavefunction, which can be written

$$|\Psi_{\text{CCD}}\rangle \approx |\phi_1\phi_2\phi_3\phi_3\rangle + |g_{12}\phi_3\phi_4\rangle + |\phi_1\phi_2g_{34}\rangle . \quad (9.27)$$

The CID function contains all doubles excitations, *but nothing more*, while the CCD function adds those quadruples excitations that are doubles excitations on *each cluster independently*. It turns out that this gives the exponential parameterization a great advantage.

## 9.2 Excitation operators, cluster operators, existence of cluster expansion

In the previous section we motivated the exponential ansatz for the exact wavefunction,

$$|\Psi\rangle = e^{\hat{T}} |\Phi\rangle, |\Phi\rangle = |123 \cdots N\rangle. \quad (9.28)$$

The cluster operator  $\hat{T}$  is a linear combination of excitation operators:

$$\hat{T} = \hat{T}_1 + \hat{T}_2 + \cdots \quad (9.29)$$

$$\hat{T}_1 = \sum_{ia} t_i^a b_a^\dagger b_i^\dagger, \quad (9.30)$$

$$\hat{T}_2 = \frac{1}{2!^2} \sum_{ijab} t_{ij}^{ab} b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger, \quad (9.31)$$

$$\hat{T}_3 = \frac{1}{3!^2} \sum_{ijk} t_{ijk}^{abc} c_c^\dagger b_k^\dagger b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger, \quad (9.32)$$

$$\vdots \quad (9.33)$$

Here, we introduced quasiparticle operators, and we assume that the *amplitudes*  $t_{ij}^{ab}$  etc are antisymmetric in the upper and lower indices, separately.

Using our notation from the MBPT section, we write  $X$  for a general excitation index, so that

$$\hat{T} = \sum_X t_X \hat{\tau}_X, \quad (9.34)$$

where  $\hat{\tau}_X = b_a^\dagger b_i^\dagger$  for  $X = \binom{a}{i}$ , etc, is an excitation operator.

Recall the exponential series,

$$e^{\hat{T}} \equiv 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \hat{T}^k. \quad (9.35)$$

Since there are only  $N$  particles in  $|\Phi\rangle$ , we can create at most  $N$  holes. Therefore  $\hat{T}^{N+1} = 0$ , and the exponential series has a finite number of terms.

The *rank* of an excitation operator or a cluster operator is the number of holes or particles created. Thus, the rank of  $b_b^\dagger b_j^\dagger b_a^\dagger b_i^\dagger$  is 2, the rank of  $\hat{T}_n$  is  $n$ , etc.

It is important to note that

- The product of two excitation operators/cluster operators is again an excitation operator/cluster operator.
- All excitation operators/cluster operators commute.
- The rank of a product is the sum of the ranks of the factors, for example,

$$\hat{T}_n \hat{T}_m = \hat{C}_{n+m}$$

- The excitation operators are *nilpotent*,

$$\hat{\tau}_X^2 = 0.$$

since the product contains two of each  $b_p^\dagger$ 's.



- $\hat{\tau}_X^\dagger$  is a *deexcitation operator* that destroys holes and particles. Thus,

$$\langle \Phi | \hat{\tau}_X = (\hat{\tau}_X | \Phi \rangle)^\dagger \equiv 0.$$

**Exercise 9.3.** Verify/prove the above points.  $\triangle$

Since all cluster operators commute, we have

$$e^{\hat{T}+\hat{T}'} = e^{\hat{T}} e^{\hat{T}'}. \quad (9.36)$$

We now prove a fundamental result in CC theory:

**Theorem 9.1** (Existence of cluster expansion). *Suppose  $\langle \Phi | \Psi \rangle = 0$ . Then there exists a (unique) cluster operator  $\hat{T}$  and a (unique) cluster operator  $\hat{A}$  such that*

$$|\Psi\rangle = (1 + \hat{A}) |\Phi\rangle = e^{\hat{T}} |\Phi\rangle. \quad (9.37)$$

*Proof.* First, we prove the linear existence part. We can expand  $|\Psi\rangle$  in the Slater determinant basis, which gives

$$|\Psi\rangle = A_0 |\Phi\rangle + \sum_X A_X |\Phi_X\rangle. \quad (9.38)$$

Since  $\langle \Phi | \Psi \rangle = 1$ , we have  $A_0 = 1$ . Since  $|\Phi_X\rangle = \hat{\tau}_X |\Phi\rangle$ , we see that

$$|\Psi\rangle = |\Phi\rangle + \sum_X A_X \hat{\tau}_X |\Phi\rangle. \quad (9.39)$$

We therefore define  $\hat{A} = \sum_X A_X \hat{\tau}_X$ , and get  $|\Psi\rangle = (1 + \hat{A}) |\Phi\rangle$ .

We now turn to the exponential part. Since we already have  $\hat{A}$  we can use this to define  $\hat{T}$ . If we can find  $\hat{T}$  such that  $1 + \hat{A} = e^{\hat{T}}$  we are finished. We split  $\hat{A}$  into singles, doubles, etc, and write

$$\hat{A} = \hat{A}_1 + \hat{A}_2 + \dots. \quad (9.40)$$

Next, we write out the exponential,

$$e^{\hat{T}} \equiv 1 + \sum_{k=1}^{\infty} \frac{1}{k!} \hat{T}^k = 1 + \hat{T} + \frac{1}{2} \hat{T}^2 + \dots. \quad (9.41)$$

Clearly,  $e^{\hat{T}} = 1 + \hat{A}$  if they have the same singles, doubles, etc, parts. We therefore write out the exponential, inserting  $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_N$ , expand each power and group terms of equal rank together:

$$\begin{aligned} e^{\hat{T}} &= 1 + \hat{T} + \frac{1}{2} \hat{T}^2 + \frac{1}{6} \hat{T}^3 + \dots \\ &= 1 + (\hat{T}_1 + \hat{T}_2 + \dots) + \frac{1}{2} (\hat{T}_1 + \hat{T}_2 + \dots)^2 + \frac{1}{6} (\hat{T}_1 + \hat{T}_2 + \dots)^3 + \frac{1}{24} (\hat{T}_1 + \hat{T}_2 + \dots)^4 \dots \\ &= 1 + \hat{T}_1 + (\hat{T}_2 + \frac{1}{2} \hat{T}_1^2) + (\hat{T}_3 + \frac{1}{6} \hat{T}_1^3 + \frac{2}{2} \hat{T}_1 \hat{T}_2) + (\hat{T}_4 + \frac{1}{24} \hat{T}_1^4 + \frac{1}{2} \hat{T}_2^2 + \frac{3}{6} \hat{T}_1^2 \hat{T}_2 + \frac{2}{2} \hat{T}_1 \hat{T}_3) + \dots \end{aligned} \quad (9.42)$$

In the final equation, we have grouped terms of equal rank together. In the rank  $n$  group, there is *only one term* that comes from  $\hat{T}_n$ , the first term in each parenthesis. All other terms are from lower rank. This is true also for the terms we have not listed.

The above expansion must be equal to  $1 + \hat{A}_1 + \hat{A}_2 + \dots$ . Therefore, *each parenthesis is equal to  $\hat{A}_n$* .

$$\hat{A}_1 = \hat{T}_1 \quad (9.43)$$

$$\hat{A}_2 = \hat{T}_2 + \frac{1}{2}\hat{T}_1^2 \quad (9.44)$$

$$\hat{A}_3 = \hat{T}_3 + \frac{1}{6}\hat{T}_1^3 + \frac{2}{2}\hat{T}_1\hat{T}_2 \quad (9.45)$$

$$\hat{A}_4 = \hat{T}_4 + \frac{1}{24}\hat{T}_1^4 + \frac{1}{2}\hat{T}_2^2 + \frac{3}{6}\hat{T}_1^2\hat{T}_2 + \frac{2}{2}\hat{T}_1\hat{T}_3 \quad (9.46)$$

Starting with rank 1, we see that  $\hat{T}_1 = \hat{A}_1$ . Plugging this into the second rank, we find

$$\hat{T}_2 = \hat{A}_2 - \frac{1}{2}\hat{T}_1^2. \quad (9.47)$$

We can similarly solve for  $\hat{T}_3, \hat{T}_4$ , recursively. Thus, given  $|\Psi\rangle$ , we can find  $\hat{A}$ , which gives  $\hat{T}$ , which exists and is unique. QED.  $\square$

Ok, we have established that the cluster expansion exists, and that it is a *nonlinear reparameterization* of the linear CI parameterization  $|\Psi\rangle = (1 + \hat{A})|\Phi\rangle$ . This parameterization has the property, that if we *truncate* the expansion  $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots + \hat{T}_K$  after a finite  $K < N$ , we obtain all possible cluster substitutions of products of occupied single-particle functions in  $|\Phi\rangle$  with rank up to and including  $\hat{K}$ . Thus, such a truncation should give a good approximation to the exact wavefunction, including *more terms* than if we had truncated  $\hat{A}$  (that's configuration-interaction!) This will give rise to a sequence of more and more accurate approximations:

$$|\Psi_{CCS}\rangle = e^{\hat{T}_1} |\Phi\rangle \quad \text{coupled-cluster singles (CCS)} \quad (9.48)$$

$$|\Psi_{CCSD}\rangle = e^{\hat{T}_1 + \hat{T}_2} |\Phi\rangle \quad \text{coupled-cluster singles and doubles (CCSD)} \quad (9.49)$$

$$|\Psi_{CCSDT}\rangle = e^{\hat{T}_1 + \hat{T}_2 + \hat{T}_3} |\Phi\rangle \quad \text{coupled-cluster singles, doubles and triples (CCSDT)} \quad (9.50)$$

$$\vdots \quad (9.51)$$

The parameters of CCS are the singles amplitudes  $t_i^a$ . For CCSD, we have  $t_i^a$  and  $t_{ij}^{ab}$ . For CCSDT we need to add  $t_{ijk}^{abc}$ , and so on.

## 9.3 The coupled-cluster equations

### 9.3.1 Unlinked equations

Since  $\hat{T}$  determines  $|\Psi\rangle$ , the question is now how to turn the Schrödinger equation into an equation for  $\hat{T}$ , i.e., for the unknown amplitudes  $t = (t_X)$ .

Let us plug the cluster expansion into the Schrödinger equation,

$$\hat{H}e^{\hat{T}}|\Phi\rangle = Ee^{\hat{T}}|\Phi\rangle. \quad (9.52)$$

Projecting onto the reference  $\langle\Phi|$  and using orthogonality gives the equations

$$E = \langle\Phi|\hat{H}e^{\hat{T}}|\Phi\rangle, \quad (9.53)$$

which determines the energy as a nonlinear function of  $\hat{T}$ . Projecting Eq. (9.52) onto an excited determinant  $\langle\Phi_X|$  gives the equation

$$E\langle\Phi_X|e^{\hat{T}}|\Phi\rangle = \langle\Phi_X|\hat{H}e^{\hat{T}}|\Phi\rangle, \quad \forall X. \quad (9.54)$$

Since we have one equation for each  $X$ , generated by projection onto a complete basis, there is precisely one equation per amplitude  $t_X$ , and we may now in theory determine  $t_X$ . Equation (9.54) is referred to as the *unlinked coupled-cluster amplitude equations*. They are cumbersome to work with, since the energy  $E$  and the amplitudes must be determined simultaneously. Moreover, the nonlinear form of the right-hand-side is too complicated to work with.

### 9.3.2 Linked equations

Instead of the unlinked equations, it is standard to work with the *linked coupled-cluster amplitude equations* found as follows: The inverse operator of  $e^{\hat{T}}$  is  $e^{-\hat{T}}$ . Let us multiply the Schrödinger equation (9.52) from the left with  $e^{-\hat{T}}$ :

$$e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle = E|\Phi\rangle. \quad (9.55)$$

Looking at Eq. (9.55), we see that it represents a *similarity transformation* of the Schrödinger equation. The similarity transformation makes  $|\Phi\rangle$  an eigenfunction. Let us project from the left with  $\langle\Phi|$ :

$$E = \langle\Phi|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle = \langle\Phi|\hat{H}e^{\hat{T}}|\Phi\rangle. \quad (9.56)$$

This is the same as Eq. (9.53). We used that  $\langle\Phi|\hat{T} \equiv 0$ . Projecting with  $\langle\Phi_X|$  instead gives

$$0 = \langle\Phi_X|e^{-\hat{T}}\hat{H}e^{\hat{T}}|\Phi\rangle, \quad \forall X. \quad (9.57)$$

Equation (9.56) is called the *CC energy equation* since it gives the energy eigenvalue as a function of  $\hat{T}$ . For all  $X$ , Equation (9.57) constitutes the *CC amplitude equations*. Since we have as many equations as we have unknowns  $t_X$ , the amplitude equations can be used to find  $t_X$ .

The linked form of the amplitude equations are of course equivalent to the unlinked ones: we only multiplied the Schrödinger equations with an invertible operator before projecting. But this multiplication brings in a similarity transformation of  $\hat{H}$  which is very convenient to work with.

We assume from now on a Hamiltonian on the form

$$\hat{H} = \hat{H}_0 + \hat{W} \quad (9.58)$$

as before. In particular,  $\hat{H}_0 = \sum_{pq} h_q^p c_p^\dagger c_q$ . We assume a given SPF basis has been assigned. We note that if we write

$$\hat{H} = \hat{H}_N + E_{\text{ref}}, \quad E_{\text{ref}} = \langle\Phi|\hat{H}|\Phi\rangle, \quad (9.59)$$

then  $H_N$  is normal-ordered,

$$\hat{H}_N = \{\hat{F}\} + \{\hat{W}\}. \quad (9.60)$$

See the section on normal-ordering of operators. We have

$$\{\hat{F}\} = \sum_{pq} f_q^p \{c_p^\dagger c_q\}, \quad \{\hat{W}\} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \{c_p^\dagger c_q^\dagger c_s c_r\}, \quad (9.61)$$

where

$$f_q^p = h_q^p + \sum_j w_{qj}^{pj}. \quad (9.62)$$

If the SPF basis happens to be the Hartree–Fock basis, then

$$E_{\text{ref}} = E_{\text{HF}}, \quad \text{and} \quad f_a^i = f_a^i = 0, \quad (9.63)$$

the latter by Brillouin’s Theorem. We will not assume a HF basis in the following except where noted explicitly.

Since a constant term in the Hamiltonian does not affect the amplitude equations (why?), we can write the CC equations as:

$$E = E_{\text{ref}} + \langle \Phi | \hat{H}_N e^{\hat{T}} | \Phi \rangle. \quad (9.64a)$$

$$0 = \langle \Phi_X | e^{-\hat{T}} \hat{H}_N e^{\hat{T}} | \Phi \rangle, \quad \forall X. \quad (9.64b)$$

The energy shift  $E - E_{\text{ref}}$  is often called *the correlation energy*.

### 9.3.3 Wick’s Theorem and The Baker–Campbell–Hausdorff expansion

The amusing thing about the CC amplitude equation, is that it is in fact a *low-order polynomial* in  $t_X$ . We already know that the exponential expansion  $e^{\pm \hat{T}}$  has only  $N$  terms, implying that the amplitude equation can be written as a polynomial of order no higher than  $2N$ . However, it turns out that the actual order is much lower, due to *exact truncation of the Baker–Campbell–Hausdorff (BCH) expansion*. The BCH expansion is

$$e^{-T} A e^T = A + [A, T] + \frac{1}{2} [[A, T], T] + \frac{1}{3!} [[[A, T], T], T] + \cdots. \quad (9.65)$$

This is an exact expansion for any pair of matrices  $A$  and  $T$ . (It can also be generalized to infinite dimensional spaces under suitable conditions.)

We will now apply the generalized Wick’s Theorem to the commutator  $[\hat{A}, \hat{T}]$ , where  $\hat{A}$  is assumed to be normal-ordered, and where  $\hat{T}$  is an excitation operator, i.e., it only contains terms with an even number of quasiparticle creation operators  $b_p^\dagger$ , and not annihilation operators  $b_p$ . Wick’s Theorem applied to the product  $\hat{A}\hat{T}$  gives

$$\hat{A}\hat{T} = \{\hat{A}\hat{T}\} + \sum_{(1)} \overline{\{\hat{A}\hat{T}\}} + \sum_{(2)} \overline{\overline{\{\hat{A}\hat{T}\}}} + \cdots \equiv \{\hat{A}\hat{T}\} + \{\hat{A}\hat{T}\}_c \quad (9.66)$$

Here we have defined the *connected part*  $\{\hat{A}\hat{T}\}_c$  of the normal-ordered product  $\{\hat{A}\hat{T}\}$  as all the terms that contain *at least one contraction (connection) between an operator in  $\hat{T}$  and an operator in  $\hat{A}$* .

For the product  $\hat{T}\hat{A}$  we note that there can be no non-zero contractions between an operator in  $\hat{T}$  and an operator in  $\hat{A}$ , since all the operators in  $\hat{T}$  are creation operators,

$$\hat{T}\hat{A} = \{\hat{T}\hat{A}\} = \{\hat{A}\hat{T}\}. \quad (9.67)$$

In the last equality we used that there is an *even* number of operators in  $\hat{T}$ . Thus, the commutator becomes,

$$[\hat{A}, \hat{T}] = \{\hat{A}\hat{T}\}_c. \quad (9.68)$$

Iteration of this argument gives:

$$[[\hat{A}, \hat{T}], \hat{T}] = \{\{\hat{A}\hat{T}\}_c \hat{T}\}_c \quad (9.69)$$

$$[[[\hat{A}, \hat{T}], \hat{T}], \hat{T}] = \{\{\{\hat{A}\hat{T}\}_c \hat{T}\}_c \hat{T}\}_c \quad (9.70)$$

$\vdots$

The second-order commutator consists of all those terms that have at least one contraction from *each* factor  $\hat{T}$  to  $\hat{A}$ . When  $\hat{A}$  is a one-body operator, there can be at most two such connections. When  $\hat{A}$  is a two-body operator, it can be at most 4 such connections, etc. Thus, the BCH expansion for  $\hat{A} = \hat{H}_N$  terminates after 4 terms, and we write

$$e^{-\hat{T}} \hat{H}_N e^{\hat{T}} = \{\hat{H}_N e^{\hat{T}}\}_c, \quad (9.71)$$

where it is to be understood that the connectedness criterion is that for each power of  $\hat{T}$  in the expansion of the exponential, there must be at least one connection with  $\hat{T}$ . The exponential has a finite number of terms: 4 for a standard two-body Hamiltonian.

The CC energy and amplitude equations can now be written

$$E = E_{\text{ref}} + \langle \Phi | \{\hat{H}_N e^{\hat{T}}\}_c | \Phi \rangle \quad (9.72a)$$

$$0 = \langle \Phi_X | \{\hat{H}_N e^{\hat{T}}\}_c | \Phi \rangle, \quad \forall X \quad (9.72b)$$

### 9.3.4 Full CC versus truncated CC

If the cluster operator  $\hat{T} = \hat{T}_1 + \hat{T}_2 + \dots$  is not truncated, the CC amplitude and energy equations are equivalent to the exact eigenvalue problem, as we have not introduced approximations. However, if we truncate  $\hat{T}$  at, say, singles-and doubles,  $\hat{T} = \hat{T}_1 + \hat{T}_2$ , all wavefunctions in the Hilbert space cannot be reached, so we cannot be sure that we solve the exact problem. Moreover, we have to truncate the amplitude equations, too, at the same level as  $\hat{T}$ , since otherwise we would have *more* equations than unknowns!

For example, let us consider the coupled-cluster singles-and-doubles approximation, CCSD, where  $\hat{T} = \hat{T}_1 + \hat{T}_2$ . We define the function  $F_X(t)$ , where  $t = (t_X)$  is the amplitude vector as

$$F_X(t) = \langle \Phi_X | \{\hat{H}_N e^{\hat{T}}\}_c | \Phi \rangle. \quad (9.73)$$

The function  $F$  should not be confused with the Fock operator  $\hat{F}$ .

In the CCSD case, the amplitude vector is  $t = (t_S, t_D)$ , where  $t_S = (t_i^a)$  is the vector of singles amplitudes, and where  $t_D = (t_{ij}^{ab})$  is the vector of doubles amplitudes. The number of elements in  $t_S$  is  $N \times N_{\text{vir}}$ , where  $N_{\text{vir}}$  is the number of virtual SPFs. The number of independent variables in  $t_D$  is  $\frac{1}{2}N(N-1) \times \frac{1}{2}N_{\text{vir}}(N_{\text{vir}}-1)$ . Correspondingly, we consider the amplitude equations

$$F_i^a(t) = 0, \quad F_{ij}^{ab}(t) = 0. \quad (9.74)$$

The higher-order equations are not used, as they would lead to more equations than unknowns. We can write the full set of CCSD equations as

$$F(t) = (F_S(t), F_D(t)) = 0. \quad (9.75)$$

## 9.4 The CCSD amplitude equations written out

A major task is now to derive the explicit forms of  $F_i^a(t)$  and  $F_{ij}^{ab}(t)$ , in a form which is programmable. To this end, we assume

$$\hat{H}_N = \hat{H} - E_{\text{ref}} = \{\hat{F}\} + \{\hat{W}\}, \quad (9.76)$$

where  $E_{\text{ref}} = \langle \Phi | \hat{H} | \Phi \rangle$ . If our basis is the Hartree–Fock basis, then  $E_{\text{ref}} = E_{\text{HF}}$ . Furthermore, we have

$$\{\hat{F}\} = \sum_{pq} f_q^p \{c_p^\dagger c_q\}, \quad \{\hat{W}\} = \frac{1}{4} \sum_{pqrs} w_{rs}^{pq} \{c_p^\dagger c_q^\dagger c_s c_r\}, \quad (9.77)$$

which are normal-ordered. In the case where the basis is the Hartree–Fock basis, we have  $f_a^i = f_a^i$  by Brillouin’s Theorem.

The full form of the CCSD equations can be found in Crawford and Schaefer. Note that they use the notation  $\langle pq || rs \rangle$  for the matrix element  $w_{rs}^{pq}$ , and that they use  $f_{pq}$  rather than  $f_q^p$  for the Fock matrix elements.

The derivation of the CCSD amplitude equations will be written here in the future. For now, use Crawford and Schaefer as a source.

### 9.4.1 A variational CC theory?

We have not tried the obvious first choice for an approximate solution to the Schrödinger equation: for The variational principle. Plugging  $\hat{T} = \sum_X t_X \hat{\tau}_X$  into the VP gives us a functional to minimize:

$$\mathcal{E}(\{t_x\}) = \frac{\langle \Phi | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} | \Phi \rangle}{\langle \Phi | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi \rangle}. \quad (9.78)$$

This is a complicated-looking object. Indeed, if we try to evaluate the functional, we quickly realize that the cluster expansion has made things very complicated for us. Historically, this *variational coupled-cluster method* has been tried but found too complicated to be practical compared to the other non-variational approach.

## 9.5 Numerical solution of the amplitude equations

We use CCSD as an example throughout this section.

How would we solve the amplitude equations numerically? These are very complicated, and expensive, so we would like a *fast* method that introduces little overhead.

As a nonlinear equation  $F(t) = 0$ , it is tempting to suggest Newton's Method, i.e., iterating

$$t^{(k+1)} = t^{(k)} - J(t^{(k)})^{-1} F(t^{(k)}), \quad (9.79)$$

where  $J(t) = F'(t)$  is the Jacobian – it is a matrix of similar size as the CISD matrix, and it needs to be both evaluated and inverted (used as a coefficient matrix for a linear solve) in every Newton iteration. This is expensive.

A quasi-Newton scheme is usually preferred, where we *approximate* the Jacobian at each step by something easily inverted.

The usual approach is the following: if the reference  $|\Phi\rangle$  is quite good, then  $t$  should be small-ish, and the fluctuation potential  $\{\hat{W}\}$  should be small-ish, too. Then we have

$$F_i^a(t) = \sum_c f_c^a t_i^c - \sum_k f_i^k t_k^a + G_i^a(t), \quad (9.80)$$

where  $G_i^a(t)$  has a small-ish derivative. This can be seen by inspecting the terms in  $G_i^a(t)$ , see the full singles equations elsewhere, e.g., Crawford and Schaefer Eq. (152). If we further assume a canonical HF basis,  $f_q^p = \delta_{pq}\epsilon_q$ , and

$$F_i^a(t) = (\epsilon_a - \epsilon_i)t_i^a + G_i^a(t). \quad (9.81)$$

This assumption is not necessary, but it simplifies our exposition. If the basis is not a canonical HF basis, the same steps can be used to define a quasi-Newton scheme. Then  $G_i^a$  will contain some extra terms. (Exercise.)

We now approximate  $J$  (for the singles part) by the derivative of the linear map isolated before  $G_i^a(t)$ . This is a simple, diagonal matrix which can be trivially inverted. We obtain the equation

$$t_i^a = (\epsilon_i - \epsilon_a)^{-1} G_i^a(t), \quad (9.82)$$

and simple iteration of this equation will be our quasi-Newton scheme,

$$(t_i^a)^{(k+1)} = (\epsilon_i - \epsilon_a)^{-1} G_i^a(t^{(k)}). \quad (9.83)$$

In a completely analogous manner we get the quasi-Newton iteration for the doubles equations,

$$(t_{ij}^{ab})^{(k+1)} = (\epsilon_i + \epsilon_j - \epsilon_a - \epsilon_b)^{-1} G_{ij}^{ab}(t^{(k)}), \quad (9.84)$$

where  $G_{ij}^{ab}$  is separated via the equation

$$F_{ij}^{ab}(t) = \sum_c (f_c^b t_{ij}^{ac} - f_c^a t_{ij}^{bc}) - \sum_k (f_j^k t_{jk}^{ab} - f_i^k t_{jk}^{ab}) + G_{ij}^{ab}(t) \quad (9.85)$$

Again, using the assumption of a canonical HF basis, we obtain

$$F_{ij}^{ab}(t) = (\epsilon_a + \epsilon_b - \epsilon_i - \epsilon_j)t_{ij}^{ab} + G_{ij}^{ab}(t) \quad (9.86)$$

## 9.6 The coupled-cluster Lagrangian

### 9.6.1 Expectation values

Quantum mechanics is not only about finding the ground-state. It is also about computing expectation values of observables in that state. Indeed, all physics can be derived from such expectation values.

The usual formula for the expectation value of  $\hat{A}$  in the (pure) state  $|\Psi\rangle$  is

$$\langle \hat{A} \rangle = \frac{\langle \Psi | \hat{A} | \Psi \rangle}{\langle \Psi | \Psi \rangle}. \quad (9.87)$$

We expect that the expectation value of  $\hat{H}$  reproduces the CC energy. However, it does not. Clearly,

$$E_{\text{CC}} = \langle \Phi | \hat{H} e^{\hat{T}} | \Phi \rangle \neq \langle \hat{H} \rangle = \frac{\langle \Phi | e^{\hat{T}^\dagger} \hat{H} e^{\hat{T}} | \Phi \rangle}{\langle \Phi | e^{\hat{T}^\dagger} e^{\hat{T}} | \Phi \rangle}. \quad (9.88)$$

In the untruncated CC case they *are* equal, since we then solve the exact problem. But when we truncate, we cannot get the same numbers: the CC energy is a quadratic polynomial in  $t$ , while the “conventional” expectation value contains  $t$  to  $N$ -th order in both the numerator and denominator.

Thus, the standard formula for the expectation value does not reproduce the computed value for the energy – which must be an expectation value!

The solution to this apparent paradox is that *CC is not variational*. The conventional expectation value formula is based on exact quantum mechanics. We have introduced a particular kind of approximation, and there is no *a priori* reason the two formulas for the energy should give the same answer.

### 9.6.2 The Hellmann–Feynman theorem

To solve our dilemma, we appeal to the Hellmann–Feynman theorem, which is basically a reformulation of first-order perturbation theory: If  $E[\hat{H}]$  is the ground-state energy as a function of the Hamiltonian  $\hat{H}$ , then perturbation theory tells us that

$$E[\hat{H} + s\hat{A}] = \sum_{n=0}^{\infty} s^n E^{(n)}, \quad (9.89)$$

where  $E^{(0)} = E[\hat{H}]$  is the unperturbed energy, and  $E^{(1)} = \langle \hat{A} \rangle$  is the first-order correction. Thus,

$$\langle \hat{A} \rangle = \left. \frac{d}{ds} E[\hat{H} + s\hat{A}] \right|_{s=0} \quad (9.90)$$

Equation (9.90) is the Hellmann–Feynman theorem. It is valid in exact quantum mechanics, and also for *variational approximations*. For non-variational approximations, it does not hold in general.

On the other hand, it makes sense to *define CC expectation values in the ground state* using the Hellman–Feynman theorem:

$$\langle A \rangle_{\text{CC}} \equiv \left. \frac{d}{ds} E_{\text{CC}}[\hat{H} + s\hat{A}] \right|_{s=0}. \quad (9.91)$$

Here, we have introduced the ground-state CC energy as function of the Hamiltonian.



### 9.6.3 CC expectation values I: direct approach

The CC ground-state energy approximation  $E_{\text{CC}}[\hat{H} + s\hat{A}]$  is obtained by solving the (possibly truncated) CC amplitude equations, obtaining the amplitude vector as a function  $t(s)$ . This gives

$$E_{\text{CC}}[\hat{H} + s\hat{A}] = \langle \Phi | e^{-\hat{T}(s)} (\hat{H} + s\hat{A}) e^{\hat{T}(s)} | \Phi \rangle. \quad (9.92)$$

We may introduce a power series,

$$t(s) \equiv t_* + st^{(1)} + O(s^2), \quad (9.93)$$

where  $t_*$  is the solution to the CC equations for the Hamiltonian  $\hat{H}$ . Expanding to first order in  $s$  gives

$$E_{\text{CC}}[\hat{H} + s\hat{A}] = \langle \Phi | e^{-\hat{T}_*} \hat{H} e^{\hat{T}_*} | \Phi \rangle + s \langle \Phi | e^{-\hat{T}_*} \hat{A} e^{\hat{T}_*} | \Phi \rangle + s \langle \Phi | e^{-\hat{T}_*} [\hat{H}, \hat{T}^{(1)}] e^{\hat{T}_*} | \Phi \rangle + O(s^2), \quad (9.94)$$

and thus

$$\langle \hat{A} \rangle_{\text{CC}} = \langle \Phi | e^{-\hat{T}_*} \hat{A} e^{\hat{T}_*} | \Phi \rangle + \langle \Phi | e^{-\hat{T}_*} [\hat{H}, \hat{T}^{(1)}] e^{\hat{T}_*} | \Phi \rangle. \quad (9.95)$$

The problem with this expression is twofold: (a)  $\hat{T}^{(1)}$  is unknown, and must be computed. It is the first-order correction in  $\hat{T}_*$  in response of the perturbation  $s\hat{A}$  of the Hamiltonian. (b) The unknown quantity  $\hat{T}^{(1)}$  depends on the observable  $\hat{A}$ , and thus *cannot be part of the description of the state*, since the state only can depend on the system Hamiltonian  $\hat{H}$ . Thus, the CC expectation value does not only depend on what we so far has thought of as the wavefunction  $|\Psi\rangle = e^{\hat{T}_*} |\Phi\rangle$ .

The solution to this is the CC Lagrangian formalism.

### 9.6.4 CC expectation values II: The CC Lagrangian formalism

Helgaker and Jørgensen discovered the CC Lagrangian in the 1980s. They realized that we can view the CC amplitude equations as a constrained minimization of the CC energy. For each amplitude equation  $F_X = 0$  introduce a Lagrange multiplier  $\lambda_X$ . Thus,  $\lambda = (\lambda_X)$  is an amplitude vector just like  $t = (t_X)$ . Define the Lagrangian

$$\begin{aligned} \mathcal{L}(t, \lambda) &= \langle \Phi | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle + \sum_X \lambda_X F_X(t) \\ &= \langle \Phi | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle + \sum_X \lambda_X \langle \Phi_X | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle \\ &= \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle, \quad \hat{\Lambda} \equiv \sum_X \lambda_X \hat{\tau}_X^\dagger. \end{aligned} \quad (9.96)$$

Recall that  $\hat{\tau}_X$  is a simple product of an even number of quasiparticle creation operators, e.g.,  $\hat{\tau}_i^a = b_a^\dagger b_i^\dagger$ . Correspondingly,  $\hat{\tau}_X^\dagger$  contains only destruction operators, and is a *deexcitation operator*. In particular,  $\langle \Phi | \hat{\tau}_X = \langle \Phi_X |$ .

Consider the partial derivatives of  $\mathcal{L}$ :

$$\frac{\partial}{\partial \lambda_X} \mathcal{L}(t, \lambda) = F_X(t) = \langle \Phi_X | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle, \quad (9.97)$$

$$\frac{\partial}{\partial t_X} \mathcal{L}(t, \lambda) \equiv \tilde{F}_X(t, \lambda) = \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} [\hat{H}, \hat{\tau}_X] e^{\hat{T}} | \Phi \rangle. \quad (9.98)$$

Here, we also defined the function  $\tilde{F}_X(t, \lambda)$ . Under constrained optimization, all partial derivatives must vanish at some point, say  $(t_*, \lambda_*)$ , i.e., we must have

$$F_X(t) = 0, \quad \tilde{F}_X(t, \lambda) = 0, \quad \forall X. \quad (9.99)$$

We obtain in particular that

$$\mathcal{L}(t_*, \lambda_*) = \langle \Phi | e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle = E_{\text{CC}}[\hat{H}], \quad (9.100)$$

since the constraints are precisely the CC amplitude equations  $F_X(t) = 0$ .

We have seen that the CC amplitude equations correspond to a constrained optimization of the CC energy, the constraints being precisely the amplitude equation  $F_X(t) = 0$ .

The numbers  $\lambda_X$  do not affect the CC energy, but they are instrumental for computing expectation values. They can be found after the CC amplitude equations have been solved, using the equation

$$\tilde{F}_Y(t_*, \lambda) = \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} [\hat{H}, \hat{\tau}_Y] e^{\hat{T}} | \Phi \rangle = b_Y(t_*) + \sum_X \lambda_X J_{XY}, \quad (9.101)$$

which is in fact a linear system of equations for the amplitudes  $\lambda_X$ . Here,

$$b_Y(t) \equiv \langle \Phi | e^{-\hat{T}} [\hat{H}, \hat{\tau}_Y] e^{\hat{T}} | \Phi \rangle \quad (9.102)$$

$$J_{XY}(t) = \frac{\partial F_X(t)}{\partial t_Y} = \langle \Phi_X | e^{-\hat{T}} [\hat{H}, \hat{\tau}_Y] e^{\hat{T}} | \Phi \rangle. \quad (9.103)$$

The matrix  $J$  is the Jacobian of  $F$ , which also figured in our discussion of Newton's method for the amplitude equations. Here, the matrix is used in a different way.

We can use the Hellmann–Feynman definition of the expectation value and derive an alternative to the direct approach in the previous subsection. To this end, introduce a notation for the Hamiltonian-dependence of the Lagrangian,

$$\mathcal{L}_{\hat{H}}(t, \lambda) = \langle \Phi | (1 + \hat{\Lambda}) e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle. \quad (9.104)$$

Introduce the function  $t(s)$  and  $\lambda(s)$ , that are the solutions to the Euler–Lagrange equations of  $\mathcal{L}_{\hat{H}+s\hat{A}}$ . That is,  $t(s)$  is as in the previous subsection, but  $\lambda(s)$  is new, and is given as the solution to  $\lambda(s)J + b = 0$ , where  $J$  and  $b$  depend on  $s$  through the Hamiltonian and  $t(s)$ .

By a similar argument as before, we introduce Taylor expansions of  $t(s)$  and  $\lambda(s)$  and compute the first-order Taylor expansion of  $E_{\text{CC}}[\hat{H} + s\hat{A}] = \mathcal{L}_{\hat{H}+s\hat{A}}(t(s), \lambda(s))$ ,

$$\begin{aligned} \mathcal{L}_{\hat{H}+s\hat{A}}(t(s), \lambda(s)) &= \mathcal{L}_{\hat{H}}(t(s), \lambda(s)) + \mathcal{L}_{s\hat{A}}(t(s), \lambda(s)) \\ &= \mathcal{L}_{\hat{H}}(t_* + st^{(1)} + \dots, \lambda_* + s\lambda^{(1)}) + s\mathcal{L}_{\hat{A}}(t_* + st^{(1)} + \dots, \lambda_* + s\lambda^{(1)}) \\ &= \mathcal{L}_{\hat{H}}(t_*, \lambda_*) + s \sum_X \left( \frac{\partial \mathcal{L}_{\hat{H}}(t_*, \lambda_*)}{\partial t_X} t_X^{(1)} + \frac{\partial \mathcal{L}_{\hat{H}}(t_*, \lambda_*)}{\partial \lambda_X} \lambda_X^{(1)} \right) + s\mathcal{L}_{\hat{A}}(t_*, \lambda_*) + O(s^2) \\ &= E_{\text{CC}}[\hat{H}] + s\mathcal{L}_{\hat{A}}(t_*, \lambda_*) + O(s^2). \end{aligned} \quad (9.105)$$

Here, we used that the partial derivatives of the Lagrangian vanish at  $(t_*, \lambda_*)$ . We obtain for the first-order term the value of the expectation value:

$$\langle \hat{A} \rangle_{\text{CC}} = \langle \Phi | (1 + \hat{\Lambda}_*) e^{-\hat{T}_*} \hat{A} e^{\hat{T}_*} | \Phi \rangle. \quad (9.106)$$

Note here, that the responses  $\lambda^{(1)}$  and  $t^{(1)}$  do not enter the expectation value – the expectation value only depends on  $t_*$  and  $\lambda_*$ , the ground-state amplitudes. Thus, we may say that these two amplitude vectors *together* describe the quantum state. As we saw earlier, the vector  $t$  is not enough.

We may even define *two wavefunctions*,

$$|\Psi\rangle = e^{\hat{T}_*} |\Phi\rangle, \quad \langle \tilde{\Psi}| = \langle \Phi | (1 + \hat{\Lambda}_*) e^{-\hat{T}_*}. \quad (9.107)$$

These satisfy  $\langle \tilde{\Psi} | \Psi \rangle = 1$ , and we have

$$\langle \hat{A} \rangle_{CC} = \langle \tilde{\Psi} | \hat{A} | \Psi \rangle . \quad (9.108)$$

This is a quite pleasing result that should be compared with the conventional variational expectation value.

## 9.7 The Equation-of-motion CC method for excited states

We have so far only talked about computing the ground-state energy using the CC method. When the solution  $\hat{T}$  has been found of the CC equations (we drop the star from now on), we have the similarity transformed Hamiltonian

$$\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}. \quad (9.109)$$

This operator has the same eigenvalues as the original Hamiltonian. If we consider the diagonalization of this operator, we will obtain the ground-state energy and wavefunctions and also excited states, regardless of whether  $\hat{T}$  is truncated to a CCSD approximation or not.

The equation-of-motion CC method (EOM-CC) is an approximation scheme for excited states, and can be defined as follows: Suppose  $\hat{T}$  is truncated at a specific level, say CCSD. Form the corresponding truncated CI matrix (here, CISD) and diagonalize it to obtain approximate eigenvalues.

The operator  $\bar{H}$  is non-Hermitian, so we must consider both the left- *and* the right eigenvalue problem, i.e.,

$$\bar{H} |R\rangle = E |R\rangle \quad (9.110)$$

$$\langle L| \bar{H} = E \langle L|. \quad (9.111)$$

In a CISD approximation, we have

$$|R\rangle = r_0 |\Phi\rangle + \sum_X r_X |\Phi_X\rangle = \hat{R} |\Phi\rangle, \quad (9.112)$$

where the sum runs over singles and doubles, and where

$$\hat{R} = r_0 \hat{1} + \sum_X r_X \hat{\tau}_X \quad (9.113)$$

is an excitation operator plus a multiplum of the unit operator  $\hat{1}$ . For the left vector, we obtain correspondingly

$$\langle L| = l_0 \langle \Phi| + \sum_X l_X \langle \Phi_X| = \langle \Phi| \hat{L}, \quad (9.114)$$

with

$$\hat{L} = l_0 \hat{1} + \sum_X l_X \hat{\tau}_X^\dagger, \quad (9.115)$$

a deexcitation operator like  $\hat{\Lambda}$  except for the term  $l_0 \hat{1}$ .

It is useful to consider the following functional, which can be thought of as an expectation value functional for the energy:

$$\mathcal{E}(r, l) = \frac{\langle L| \bar{H} |R\rangle}{\langle L| R\rangle} = \frac{\sum'_{XY} l_X \langle \Phi_X| \bar{H} | \Phi_Y\rangle r_Y}{\sum'_X l_X r_X}. \quad (9.116)$$

Here, we have introduced amplitude vectors  $l$  and  $r$ . On the far right we have expanded the expression as a fraction of a vector-matrix-vector product and a vector dot product. The prime on the sum indicates that the reference  $X = Y = 0$  is included, i.e.,  $|\Phi_0\rangle \equiv |\Phi\rangle$ .

It is straightforward (exercise) to compute the partial derivatives of  $\mathcal{E}(l, r)$  and find that all partial derivatives vanish, i.e.,

$$\frac{\partial \mathcal{E}(r, l)}{\partial l_X} = 0 \quad \forall X \quad (9.117)$$

$$\frac{\partial \mathcal{E}(r, l)}{\partial r_X} = 0 \quad \forall X \quad (9.118)$$

if and only if (in terms of CISD matrices and vectors)

$$\bar{H} |R\rangle = E |R\rangle, \quad \langle L | \bar{H} = E \langle L |, \quad \langle L | R \rangle \neq 0, \quad (9.119)$$

with the eigenvalue given by

$$E = \mathcal{E}(r, l). \quad (9.120)$$

Thus the energies are obtained as stationary values for the functional  $\mathcal{E}$ , and the stationary points are the left- and right eigenvectors.

We now consider the block form of the CISD matrix of  $\bar{H}$  (exercise):

$$\bar{H} = \begin{bmatrix} E_{CC} & \bar{H}_{0S} & \bar{H}_{0D} \\ 0 & \bar{H}_{SS} & \bar{H}_{SD} \\ 0 & \bar{H}_{DS} & \bar{H}_{DD} \end{bmatrix} \quad (9.121)$$

The zero blocks come from the CCSD amplitude equations. We can see that the vector  $|R_0\rangle = |\Phi\rangle$  is a solution to the right eigenvalue problem with eigenvalue  $E = E_{CC}$ , i.e.,

$$\hat{R}_0 = \hat{1}. \quad (9.122)$$

We now seek the corresponding left eigenvector  $\langle L_0 |$ . The normalization of  $\langle L_0 |$  is arbitrary, so we require  $\langle L_0 | R_0 \rangle = 1$ , which implies that  $(l_0)_0 = 1$ , i.e., we may set  $\hat{L}_0 = \hat{1} + \tilde{L}$ , for some deexcitation cluster operator  $\tilde{L}$ . We now note that

$$\mathcal{E}(l_0, r_0) = \langle \Phi | (\hat{1} + \tilde{L}) e^{-\hat{T}} \hat{H} e^{\hat{T}} | \Phi \rangle = \mathcal{L}(\tilde{L}, t). \quad (9.123)$$

Thus, the derivative of  $\mathcal{E}$  with respect to  $l_0$  vanishes if and only if  $\tilde{L} = \lambda$  from the ground-state problem!

Thus, we have found the CC ground state as the ground state of the EOM-CC matrix.

The other eigenvalues and eigenvectors are approximations to excited states. Computationally, it is similar to a CISD calculation, except that the matrix is non-Hermitian.

### 9.7.1 Exercises

**Exercise 9.4.** Show that the partial derivatives of  $\mathcal{E}(r, l)$  vanish simultaneously if and only if the right and left eigenvalue equations are satisfied simultaneously, as described in the text.  $\triangle$

**Exercise 9.5.** Show that the CISD matrix of  $\bar{H} = e^{-\hat{T}} \hat{H} e^{\hat{T}}$  has the block structure given in Eq. (9.121). Recall that  $\hat{T}$  is the CCSD solution.

Do the same for CCSDT.  $\triangle$

## 9.8 Coupled-cluster diagrams

### 9.8.1 Introduction

Even Wick's Theorem can be hard to apply to generate the various terms in the CC amplitude equations. Diagrammatic analysis based on Feynman-like diagrams is a popular alternative. It is a visual approach to generating algebraic expressions, and can be used also for perturbation theory and other topics such as Green's functions.

Many different formulations of diagrammatic manybody theory exist. We here follow Crawford and Schaefer, using so-called *anti-symmetrized Goldstone diagrams*.

Diagrams can be used to represent virtually any kind of object: wavefunctions, operators, operator-wavefunction products, matrix elements, and so on. The book by Shavitt and Bartlett (MBPT and CC theory) is an excellent additional sourcebook. Also, the book by Harris, Freeman and Monkhorst is a solid reference on many aspects.

We will not introduce diagrams in a mathematically rigorous manner. This will take too much effort. Instead, we introduce various aspects of diagrams as we go along, illustrating these with comparing with the treatment from Wick's Theorem.

We will only scratch the surface of diagrammatic CC theory, so that the student has a working knowledge and a starting point when she/he needs to deal with diagrams later in her/her career.

### 9.8.2 Diagram elements

States.

Operators: onebody, twobody, excitation operators.

Excitation levels.

### 9.8.3 CC energy equation

### 9.8.4 CCSD amplitude equations

### 9.8.5 Exercises

## Chapter 10

# Multireference systems

### 10.1 Single-reference vs. multireference

A single-reference (SR) wavefunction is a wavefunction that can be well-described with a single Slater determinant  $|\Phi\rangle$ . A single-reference *system* is a system whose wavefunction (usually the ground-state wavefunction) is SR. Thus, a single-reference wavefunction can be written, in intermediate normalization,

$$|\Psi\rangle = |\Phi\rangle + \epsilon |\Theta\rangle, \quad \|\Theta\| = 1, \quad (10.1)$$

where  $\epsilon$  is a “small” number. How small? In chemistry, one has found that  $\epsilon \lesssim 0.2$  gives a good characterization, in the sense that the methods (HF, CC, ...) that are designed to work for SR systems actually do work, i.e., are stable and gives accurate enough results. These are all relative terms, of course.

Multireference (MR) wavefunctions, on the other hand, can *not* be well approximated by a single determinant, i.e., one needs “multiple references” to describe it.

In chemistry, MR systems are a *problem* because it can be very hard to determine whether a system is MR or SR, and hence which method that would work well for it. Moreover, most SR methods are *black box* and codes can be run by practitioners with little or no knowledge of how the method works. MR methods, on the other hand, are much more difficult to apply, since one needs to determine the structure of the multiple references. Currently, there are no simple ways to do so, and it requires a lot of experience and insight into the quantum mechanics of the system.

We note that while the SR definition is basis independent, in most situations one cannot actually search for a basis so that  $|\Psi\rangle$  becomes SR relative to this basis. Instead, the basis is *given*, and the single-reference nature of the wavefunction is defined relative to this basis. Typically, the basis is a Hartree–Fock basis, and it may be restricted or unrestricted.

Consider for example the dissociation of the Hydrogen molecule  $H_2$ , a two-electron system. The RHF wavefunction is on the form

$$|\Phi_{\text{RHF}}\rangle = |\varphi\bar{\varphi}\rangle = c_{1,\uparrow}^\dagger c_{1,\downarrow}^\dagger |-\rangle, \quad (10.2)$$

where  $\varphi(\vec{r})$  is a single doubly occupied orbital. As the molecule dissociates, the function  $\varphi(\vec{r})$  is *delocalized* so that the probability density is symmetric and centered on both atoms, which become far apart. Contrast this with the UHF wavefunction,

$$|\Phi_{\text{UHF}}\rangle = |\varphi_L\bar{\varphi}_R\rangle = c_{L,\uparrow}^\dagger c_{R,\downarrow}^\dagger |-\rangle, \quad (10.3)$$

where  $\varphi_L$  and  $\varphi_R$  are, in the dissociation limit, localized functions on each atom, obtaining much lower energy.

Both the UHF and the RHF calculations yield a spin-orbital basis. These bases are different, giving a different set of Slater determinant bases. But since the underlying *atomic basis is the same*, the total set of many-electron wavefunctions is also the same. Yet, the UHF wavefunction is much closer to the exact ground state than the RHF wavefunction, which means that the ground-state wavefunction is single-reference with respect to the UHF basis, but multireference with respect to the RHF basis.

Thus, in the RHF basis, we have a multireference nature of the ground-state, while for the UHF basis we have a single-reference nature of the wavefunction.

There is another way which this system is multireference: The Hamiltonian commutes with both the spin projection and the total spin of the two electrons. Thus, we may simultaneously diagonalize all three operators. If we insist on the computed wavefunction to be an eigenfunction of total spin, we have to modify the UHF wavefunction, as this is not a total spin eigenfunction. The UHF wavefunction is, in coordinate basis,

$$\Phi_{\text{UHF}}(1, 2) = \frac{1}{\sqrt{2}} [\varphi_L(1)\varphi_R(2)\chi_{\uparrow}(1)\chi_{\downarrow}(2) - \varphi_R(1)\varphi_L(2)\chi_{\downarrow}(1)\chi_{\uparrow}(2)] \quad (10.4)$$

This wavefunction does not have a definite symmetry in the spin part. The properly symmetrized singlet wavefunction is

$$\begin{aligned} \Phi_{\text{singlet}}(1, 2) &= \frac{1}{\sqrt{2}} [\varphi_L(1)\varphi_R(2) + \varphi_R(1)\varphi_L(2)] \cdot \frac{1}{\sqrt{2}} [\chi_{\uparrow}(1)\chi_{\downarrow}(2) - \chi_{\downarrow}(1)\chi_{\uparrow}(2)] \\ &= \frac{1}{\sqrt{2}} [\varphi_L(1)\varphi_R(2)\chi_{\uparrow}(1)\chi_{\downarrow}(2) - \varphi_R(1)\varphi_L(2)\chi_{\downarrow}(1)\chi_{\uparrow}(2)] \\ &\quad - \frac{1}{\sqrt{2}} [\varphi_L(1)\varphi_R(2)\chi_{\downarrow}(1)\chi_{\uparrow}(2) - \varphi_R(1)\varphi_L(2)\chi_{\uparrow}(1)\chi_{\downarrow}(2)]. \end{aligned} \quad (10.5)$$

That is,

$$|\Phi_{\text{singlet}}\rangle = \frac{1}{\sqrt{2}} [c_{L,\uparrow}^{\dagger} c_{R,\downarrow}^{\dagger} - c_{L,\downarrow}^{\dagger} c_{R,\uparrow}^{\dagger}] |-\rangle, \quad (10.6)$$

a linear combination of *two* UHF determinants where the spins are exchanged. The energy expectation value of these two determinants are identical since the Hamiltonian does not include spin at all.

It is an empirical fact, that the ground-state of most molecules close to the equilibrium geometry is of SR character, even for a RHF basis. Excited states may well be of MR character (which makes the EOMCC method a particularly nice approach – the wavefunction is explicitly written as a MR wavefunction).



## 10.2 Multiconfiguration Hartree–Fock

In this course, we will consider one genuine MR method: multiconfiguration Hartree–Fock (MCHF), often called multiconfiguration self-consistent field (MCSCF) in chemistry. A *configuration* refers to so-called *configuration-state functions* (CSFs). These are linear combinations of Slater determinants so that we obtain eigenfunctions of total spin as well as the spin-projection.

We will not work with CSFs, but instead formulate a general MCHF in terms of general single-particle functions, and not spin-orbitals. If we on the other hand specialize to spin-orbitals, we have a restricted variant of MCHF and an unrestricted variant, depending on whether we force the spatial parts of each spin-orbital be independent of spin or not, similar to RHF versus UHF. We mention this again later.

### 10.2.1 Wavefunction ansatz

The configuration-interaction wavefunction is able to describe multireference wavefunctions, since we are not limited to the HF reference having a dominating contribution. However, for a multireference system, *the HF starting point is a bad one to begin with*, in the sense that the whole idea of HF – approximating the system by one single determinant – breaks down. Thus, it may well be that the number of determinants included in the CI wavefunction is unnecessarily large. (Moreover, CISD is not size-extensive in comparison to CCSD, for example.) A better approach would be to optimize the SPFs together with the coefficients  $A_\mu$ . This is the heart of the MCHF ansatz.

The MCHF wavefunction is a linear combination of Slater determinants,

$$|\Psi\rangle = \sum_{\mu \in \mathcal{M}} A_\mu |\Phi_\mu\rangle, \quad (10.7)$$

where  $\mu$  enumerates a set  $\mathcal{M}$  of Slater determinants built from a relatively small set of SPFs  $\phi_i$ ,  $i = 1, 2, \dots, M$ . In this section, we are going to assume that the set  $\mathcal{M}$  is the set of *all* Slater determinants, i.e., a full CI wavefunction containing  $\binom{M}{N}$  terms.

The energy expectation value  $E_{\text{MCHF}}$  is now to be optimized both with respect to the amplitudes  $A_\mu$  and the SPFs  $\phi_i$ . In general, the set  $\{\phi_i\}_{i=1}^M$  is not complete ( $M$  is relatively small), but it may be augmented to a complete basis by adding  $K - M$  *virtual* functions  $\phi_a$ , assuming that the one-particle Hilbert space is  $K$  dimensional.<sup>1</sup> Let therefore  $\phi_i$  enumerate the  $M$  unknown SPFs, which we call *active*, and  $\phi_a$  enumerate the  $K - M$  virtual ones, much like in Hartree–Fock theory, except now  $K > N$  in general. Moreover, not all active SPFs are occupied at all times. We stress that the virtual functions are not part of the wavefunction. We let  $p, q$ , etc., enumerate the general SPFs, both active and virtual. See Figure 10.1 for an illustration.

We introduce the vector  $A$  of all the amplitudes  $A_\mu$ , and a matrix-like notation  $\phi$  for the complete basis of SPFs,

$$\phi = (\phi^{\text{act}}, \phi^{\text{vir}}) = (\phi_1, \phi_2, \dots, \phi_M, \phi_{M+1}, \dots, \phi_K). \quad (10.8)$$

The MCHF wavefunction is now parameterized by  $A$  and  $\phi^{\text{act}}$ ,

$$|\Psi_{\text{MCHF}}\rangle = |\Psi(A, \phi^{\text{act}})\rangle.$$

One can imagine each  $\phi_p$  expanded in a computational basis (atomic basis in chemistry), and  $\phi$  then becomes a matrix where each column is a basis expansion. Orthonormality of the  $\phi_p$  is expressed symbolically as

$$\phi^H \phi = I_K, \quad (10.9)$$

<sup>1</sup>The true problem has an infinite dimensional one-body space, but in every computation we have to introduce a finite-dimensional basis with  $K$  elements. We could formulate the present theory with infinitely many virtual functions, but we will not do that for simplicity.

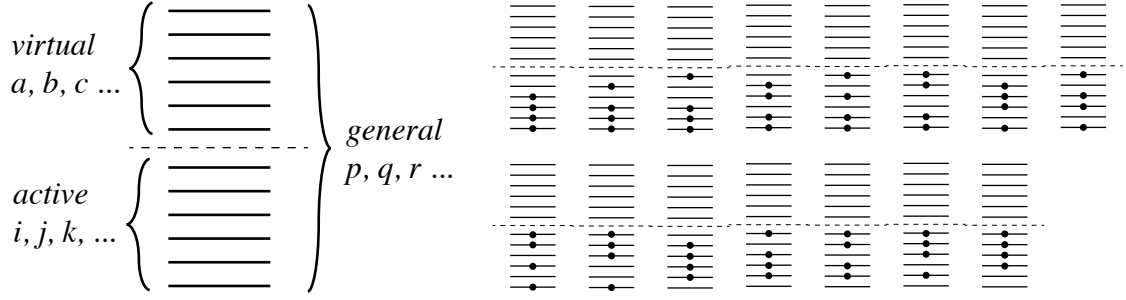


Figure 10.1: SPF structure of the MCHF wavefunction. Active and virtual SPFs are indicated, and the set of 15 determinants generated by 4 fermions in 6 active SPFs are illustrated. The virtual SPFs are never occupied.

the identity matrix. This is so because the matrix element  $(\phi^H \phi)_{pq}$  is the inner product of the  $p$ th row of  $\phi^H$  and the  $q$ th column of  $\phi$ , i.e.,  $\langle \phi_p |$  and  $|\phi_q \rangle$ , respectively. (This is consistent with expressing the  $\phi_p$  in an orthonormal and fixed computational basis.) Orthonormality of the active SPFs alone can be expressed as

$$(\phi^{\text{act}})^H \phi^{\text{act}} = I_M. \quad (10.10)$$

Orthogonality of the active and virtual spaces is expressed as

$$(\phi^{\text{vir}})^H \phi^{\text{act}} = 0_{K-M, M}. \quad (10.11)$$

### 10.2.2 Unitary transformations of SPF bases

A common technique for describing the SPF basis as a variable is to introduce a “reference” basis  $\hat{\phi} = (\hat{\phi}_1, \dots, \hat{\phi}_M, \hat{\phi}_{M+1}, \dots)$  which is considered fixed, and obtaining  $\phi$  via a unitary transformation matrix  $U$ ,

$$\phi_p = \sum_q \hat{\phi}_q U_{qp}. \quad (10.12)$$

We partition  $U = (U^{\text{occ}}, U^{\text{vir}})$ , where  $U^{\text{occ}}$  is the  $M$  first columns, i.e.,  $\phi_i = \sum_q \hat{\phi}_q U_{qi}^{\text{occ}}$ .

The choice of  $\hat{\phi}_p$  is completely arbitrary, and we may change it at will, only remembering that the definition of  $S$  will also have to change.

Any unitary matrix can be written as an exponential  $U = e^S$  with,  $S = -S^H$ . From exercise 2.22, we have that the creation operators  $c_p^\dagger = c^\dagger(\phi_p)$  are related to  $\hat{c}_p^\dagger = c^\dagger(\hat{\phi}_p)$  via the expression

$$c_p^\dagger = e^{\hat{S}} \hat{c}_p^\dagger e^{-\hat{S}}, \quad \hat{S} = \sum_{pq} S_{qp} \hat{c}_q^\dagger \hat{c}_p. \quad (10.13)$$

Moreover,

$$c_p = e^{\hat{S}} \hat{c}_p e^{-\hat{S}}. \quad (10.14)$$

Inverting the relationships gives  $\hat{c}_p^\dagger = e^{-\hat{S}} c_p^\dagger e^{\hat{S}}$  and  $\hat{c}_p = e^{-\hat{S}} c_p e^{\hat{S}}$ . For any Slater determinant built from the  $\phi_p$ ,

$$|\Phi_\mu\rangle = |\phi_{i_1} \phi_{i_2} \dots \phi_{i_N}\rangle, \quad (10.15)$$

we define

$$|\hat{\Phi}_\mu\rangle = |\hat{\phi}_{i_1} \hat{\phi}_{i_2} \dots \hat{\phi}_{i_N}\rangle. \quad (10.16)$$

This gives

$$|\Phi_\mu\rangle = e^{\hat{S}} |\hat{\Phi}_\mu\rangle. \quad (10.17)$$

Thus,  $e^{\hat{S}}$  acts as a change-of-basis operator, transforming any occupancy of  $\hat{\phi}_p$  to an occupancy of  $\phi_p$  instead. The hatted Slater determinant basis is *fixed*, and we have

$$|\Psi\rangle = e^{\hat{S}} |\hat{\Psi}\rangle, \quad (10.18)$$

where

$$|\hat{\Psi}\rangle = \sum_{\mu} A_{\mu} |\hat{\Phi}_{\mu}\rangle. \quad (10.19)$$

is only a function of  $A$ , and not of the SPFs. We see that the functional dependence of  $|\Psi\rangle$  of the basis and the amplitudes is separated, and the basis dependence is described with the matrix  $S$  instead of  $\phi$  directly,

$$|\Psi(A, S)\rangle = e^{\hat{S}} |\hat{\Psi}(A)\rangle. \quad (10.20)$$

### 10.2.3 Energy expression

The energy of the MCHF wavefunction is

$$E_{\text{MCHF}} = \frac{\langle \Psi(A, \phi^{\text{act}}) | \hat{H} | \Psi(A, \phi^{\text{act}}) \rangle}{\langle \Psi(A, \phi^{\text{act}}) | \Psi(A, \phi^{\text{act}}) \rangle} \quad (10.21)$$

We now obtain *two* expressions for the expectation value. Equation (??) readily gives

$$E_{\text{MCHF}} = \frac{A^H K A}{A^H A}, \quad K_{\mu\nu} = K_{\mu\nu}(\phi^{\text{act}}) = \langle \Phi_{\mu} | \hat{H} | \Phi_{\nu} \rangle, \quad (10.22)$$

and we note that the matrix elements are functions of  $\phi^{\text{act}}$ , but not of  $A$ . Equation (10.22) is convenient when we want to compute the variations in  $E_{\text{MCHF}}$  with respect to  $A$ .

Next, we introduce the *reduced one-body density matrix*  $\rho_p^q$  via the formula

$$\rho_p^q = \frac{\langle \Psi(A, \phi^{\text{act}}) | c_p^{\dagger} c_q | \Psi(A, \phi^{\text{act}}) \rangle}{\langle \Psi(A, \phi^{\text{act}}) | \Psi(A, \phi^{\text{act}}) \rangle}. \quad (10.23)$$

The reduced density matrix is thus the expectation value of  $c_p^{\dagger} c_q$ , and it is in fact a function of  $A$  alone. To see this, we note that if  $X_1 \cdots X_n$  is a string of creation and annihilation operators, the expectation value is given by

$$\langle X_1 \cdots X_n \rangle = (A^H A)^{-1} \sum_{\mu\nu} A_{\mu} \langle \Phi_{\mu} | X_1 \cdots X_n | \Phi_{\nu} \rangle A_{\nu}. \quad (10.24)$$

The matrix elements  $\langle \Phi_{\mu} | X_1 \cdots X_n | \Phi_{\nu} \rangle$  can in principle be evaluated with Wick's Theorem, which only relies on the orthonormality of the SPFs. Thus,  $\langle X_1 \cdots X_n \rangle$  only depends on  $A$  and the fact that  $\phi^{\text{act}}$  is an orthonormal set.

Similarly, we define the reduced two-body density matrix

$$\rho_{pr}^{qs} \equiv \frac{\langle \Psi(A, S) | c_q^{\dagger} c_s^{\dagger} c_r c_p | \Psi(A, S) \rangle}{\langle \Psi(A, S) | \Psi(A, S) \rangle} \quad (10.25)$$

which is likewise a function of  $A$  alone.

We note that all matrix elements of the reduced density matrices vanish unless all indices are active, e.g.,  $\rho_a^i, \rho_b^a$ , and  $\rho_i^a$  are all zero, as is  $\rho_{ak}^{ij}, \rho_{kb}^{ia}$ , etc. We note that the two-body reduced density matrix is antisymmetric,  $\rho_{pr}^{qs} = -\rho_{pr}^{sq} = -\rho_{rp}^{qs} = \rho_{rp}^{sq}$ .

Assuming now the Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{W}$  as usual, we obtain

$$E_{\text{MCHF}} = \sum_{ij} \rho_j^i(A) h_i^j(\phi^{\text{act}}) + \frac{1}{2} \sum_{ijkl} \rho_{ik}^{jl}(A) w_{jl}^{ik}(\phi^{\text{act}}), \quad (10.26)$$

where

$$w_{jl}^{ik}(\phi^{\text{act}}) = \langle \phi_i \phi_k | \hat{w} | \phi_j \phi_l \rangle, \quad (10.27)$$

$$h_j^i(\phi^{\text{act}}) = \langle \phi_i | \hat{h} | \phi_j \rangle. \quad (10.28)$$

The two-body matrix elements are assume *not antisymmetrized*. At this point, exercise 10.2 can be recommended.

The expression (10.26) is convenient when we want to vary the SPFs in the variational principle.

### 10.2.4 Variational equations: variation of $A$

In the variational principle, we require the energy  $E_{\text{MCHF}}$  to be stationary under all variations of the variables  $(A, \phi^{\text{act}})$ . Recall that  $A$  is a vector,  $\phi^{\text{act}} = (\phi_1, \dots, \phi_M)$  are  $M$  orthonormal SPFs. We let the variational optimum be  $(A_*, \phi_*^{\text{act}})$ , and this point is found by solving the condition  $\delta E_{\text{MCHF}}(A_*, \phi_*^{\text{act}}) = 0$ .

We begin with the amplitude variations. Doing an arbitrary variation  $\delta A_\mu$  in each amplitude gives,

$$\delta E_{\text{MCHF}}(A_*, \phi_*^{\text{act}}) = \sum_{\mu} \delta A_{\mu}^* \frac{\partial E_{\text{MCHF}}(A_*, \phi_*^{\text{act}})}{\partial A_{\mu}^*} + \sum_{\mu} \delta A_{\mu} \frac{\partial E_{\text{MCHF}}(A_*, \phi_*^{\text{act}})}{\partial A_{\mu}}, \quad (10.29)$$

which we require to vanish. Treating  $A_\mu$  and  $A_\mu^*$  as independent variables, it turns out that we only need to vary, say,  $A_\mu^*$ . Computation of the partial derivative and showing that  $A_\mu^*$  variations are sufficient is an elementary exercise (Exercise 10.4), and one obtains

$$\frac{\partial E_{\text{MCHF}}(A, \phi^{\text{act}})}{\partial A_{\mu}^*} = \frac{1}{A^H A} [(KA)_{\mu} - E_{\text{MCHF}} A_{\mu}]. \quad (10.30)$$

Here,  $K = K(\phi^{\text{act}}) = [\langle \Phi_{\mu} | \hat{H} | \Phi_{\nu} \rangle]$  is the FCI matrix, and  $(KA)_{\mu} = \sum_{\nu} K_{\mu\nu} A_{\nu}$  is the  $\mu$ -th component of the matrix-vector product. Since at the variational optimum  $(A_*, \phi_*^{\text{act}})$  all partial derivatives must vanish, we have the eigenvalue equation

$$K(\phi_*^{\text{act}}) A_* = E_* A_*, \quad A_* \neq 0, \quad E_* = E_{\text{MCHF}}(A_*, \phi_*^{\text{act}}). \quad (10.31)$$

This has the form of the FCI Schrödinger equation in the active basis. However, we need to find an equation for  $\phi_*^{\text{act}}$  as well.

### 10.2.5 Variational equations: variation of $\phi^{\text{act}}$

The SPFs  $\phi = (\phi^{\text{act}}, \phi^{\text{vir}})$  are a basis for single-particle space. Thus, any variation in  $\phi_p$  can be expanded in this basis. Moreover, since we need our SPFs to be orthonormal, we have, for any other orthonormal basis  $\phi' = (\phi^{\text{act}'}, \phi^{\text{vir}'})$  the existence of a unitary  $U = e^S$  with  $S = -S^H$ , such that

$$\phi'_p = \sum_q \phi_q (e^S)_{qp}, \quad S = -S^H. \quad (10.32)$$

It is useful to consider the corresponding transformation in creation and annihilation operators. From exercise 2.22 (which should be done now if not already done), we have that the creation operators  $c_p^\dagger = c^\dagger(\phi_p)$  are related to  $d_p^\dagger = c^\dagger(\phi'_p)$  via the expression

$$d_p^\dagger = e^{\hat{S}} c_p^\dagger e^{-\hat{S}}, \quad \hat{S} = \sum_{pq} S_{qp} c_q^\dagger c_p. \quad (10.33)$$

Notably, the Slater determinants transform as

$$|\phi'_{p_1} \phi'_{p_2} \cdots \phi'_{p_N}\rangle = e^{\hat{S}} |\phi_{p_1} \phi_{p_2} \cdots \phi_{p_N}\rangle, \quad (10.34)$$

and in particular  $|\Phi'_\mu\rangle = e^{\hat{S}} |\Phi_\mu\rangle$  is the transformed Slater determinant that enter  $|\Psi(A, \phi^{\text{act}'})\rangle$ .

If we require the change in the basis to be *infinitesimal*, we see that

$$\delta\phi_p = \phi'_p - \phi_p \approx \sum_q \phi_q S_{qp}. \quad (10.35)$$

Similarly,

$$\delta c_p^\dagger = d_p^\dagger - c_p^\dagger \approx \sum_q c_q^\dagger S_{qp}, \quad (10.36)$$

etc. The matrix elements  $S_{pq}$  then parameterizes the SPF variations, similar as in HF theory.

Note that the unitary  $K \times K$  matrix  $e^S$  couples both active and virtual SPFs. The matrix  $S$  can be divided into blocks,

$$S = \begin{pmatrix} S^{\text{act}} & -\eta^H \\ \eta & S^{\text{vir}} \end{pmatrix}, \quad (10.37)$$

where  $S^{\text{act}}$  is  $M \times M$  and couples only active SPFs with other active SPFs to first order. Similarly,  $S^{\text{vir}}$  couples only virtual SPFs to first order, while  $\eta = [\eta_{ai}]$  couples virtual to active.

Before we compute the variations in the energy, we make an observation. Suppose  $S$  is such that such that  $S_i^a = S_a^i = 0$ . That is, we have the block form

$$S = \begin{pmatrix} S^{\text{act}} & 0 \\ 0 & S^{\text{vir}} \end{pmatrix}. \quad (10.38)$$

We still assume that  $S^H = -S$ . Now, the unitary matrix  $e^S$  is also on this form (exercise 10.5),

$$U = \begin{pmatrix} U^{\text{act}} & 0 \\ 0 & U^{\text{vir}} \end{pmatrix}. \quad (10.39)$$

Thus, it only transforms active SPFs into linear combinations of active SPFs, and similarly with virtual SPFs. This means that  $e^{\hat{S}}$ , the many-electron basis change operator, maps the space spanned by the active Slater determinants  $|\Phi_\mu\rangle$  into itself. Thus, the matrix  $V_{\beta\nu} = \langle\Phi_\beta|e^{\hat{S}}|\Phi_\nu\rangle$  is unitary. (If  $e^S$  maps some active SPF out of the active space,  $e^{\hat{S}}$  will map some determinant out of the determinant space, and then the matrix  $V$  cannot be unitary.)

We now change the active SPFs according to this transformation  $U = e^S$ , and also allow for some transformation  $A' = A + \delta A$ . Computing the new state, we obtain

$$\begin{aligned} |\Psi(A', \phi^{\text{act}'})\rangle &= \sum_\mu A'_\mu e^{\hat{S}} |\Phi_\mu\rangle \\ &= \sum_\mu A'_\mu \sum_\nu |\Phi_\nu\rangle \langle\Phi_\nu| e^{\hat{S}} |\Phi_\mu\rangle = \sum_\nu B_\nu e^{\hat{S}} |\Phi_\nu\rangle, \quad B_\nu = \sum_\mu V_{\nu\mu} A'_\mu. \end{aligned} \quad (10.40)$$

Since  $A'$  can be anything, we set  $A' = V^H A$ , so that  $B = VV^H A = A$ . We see then that *the total wavefunction is invariant*. Thus, we may compensate for an active-active transformation of the SPFs by a unitary transformation of the amplitudes. The matrix elements  $S_{ij}^{\text{act}}$  are redundant variation parameters, since we treat all the amplitudes as independent variables. The matrix elements  $S_{ab}^{\text{vir}}$  are also redundant, since they do not enter the wavefunction at all.

In conclusion, we may assume that

$$S = \begin{pmatrix} 0 & -\eta^H \\ \eta & 0 \end{pmatrix}, \quad (10.41)$$

where  $\eta = [\eta_{ai}]$ . That is,

$$\delta\phi_i = \sum_a \phi_a \eta_{ai}, \quad \delta\phi_a = -\sum_i \phi_i \eta_{ai}^*. \quad (10.42)$$

The energy expression (10.26) is now suitable for variation, and we obtain

$$\delta E_{\text{MCHF}}(A, \phi^{\text{act}}) = \sum_{ij} \rho_i^j \langle \delta\phi_i | \hat{h} | \phi_j \rangle + \sum_{ijkl} \rho_{ik}^{jl} \langle \delta\phi_i \phi_j | \hat{w} | \phi_k \phi_l \rangle, \quad (10.43)$$

where the factor  $\frac{1}{2}$  in front of the twobody term in the energy is cancelled by symmetry of the matrix elements. Inserting the expansion of  $\delta\phi_i$  we obtain

$$\delta E_{\text{MCHF}} = \sum_{ia} \eta_{ai}^* \left[ \sum_j \rho_i^j \langle \phi_a | \hat{h} | \phi_j \rangle + \sum_{jkl} \rho_{ik}^{jl} \langle \phi_a \phi_j | \hat{w} | \phi_k \phi_l \rangle \right], \quad (10.44)$$

This variation has to vanish for all choices of the matrix elements  $\eta_{ai}$ , so the bracket must vanish identically at the critical point. Thus, we must solve  $F_{ai}(\phi_*) = 0$ , for all  $(a, i)$ , where

$$F_{ai}(\phi) \equiv \sum_j \rho_i^j \langle \phi_a | \hat{h} | \phi_j \rangle + \sum_{jkl} \rho_{ik}^{jl} \langle \phi_a \phi_j | \hat{w} | \phi_k \phi_l \rangle. \quad (10.45)$$

When  $F_{ai} = 0$  for all  $(a, i)$  we have a stationary point  $\phi_*^{\text{act}}$  of  $E_{\text{MCHF}}$  with respect to variations in the SPFs.

We may rewrite  $F_{ai} = 0$  on a form which explicitly exhibits the dependence of  $\phi^{\text{act}}$  only – the equations must be invariant with respect to unitary transformations among the virtual SPFs, since the wavefunction does not depend on these SPFs. We multiply the equation  $F_{ai} = 0$  with  $\sum_a |\phi_a\rangle$ , and introduce the operator  $\hat{Q} = \sum_a |\phi_a\rangle \langle \phi_a|$  and obtain

$$\Omega_i(\phi^{\text{act}}) := \hat{Q} \left[ \sum_j \rho_i^j \hat{h} | \phi_j \rangle + \sum_{jkl} \rho_{ik}^{jl} \langle \cdot | \phi_j | \hat{w} | \phi_k \phi_l \rangle \right] = 0. \quad (10.46)$$

The operator  $\hat{Q}$  is the orthogonal projection onto the virtual SPF space: for any SPF  $\psi$ , we have

$$\hat{Q} |\psi\rangle = |\psi\rangle - \sum_i |\phi_i\rangle \langle \phi_i | \psi \rangle. \quad (10.47)$$

Equations (10.46) and (10.45) are two equivalent formulations of the MCHF equation for determining the active SPFs  $\phi_i$ . Introducing the operator  $\hat{Q}$  has the benefit that  $\Omega$  is explicitly a function only of the active SPFs –  $\hat{Q}$  is a projector onto the orthogonal complement of the active space, and does not depend on the basis  $\phi_a$  chosen. Thus, the last equation is formulated only in terms of the unknown SPFs. The former equation can also be used, and may be more practical.

The equation  $\Omega_i = 0$  is a *self-consistent mean-field equation* just like Hartree–Fock, except that the field set up depends on more than one determinant. Indeed, using the HF reduced density matrices, it can be seen that the equation reduces to the HF equation. See exercise 10.3.

## 10.2.6 Newton's Method for MCHF

We briefly discuss the formulation of Newton's method for the equation  $F_{ai} = 0$ . This is today a common method, at least in chemistry, for solving the MCHF mean-field equations. On the other hand, it is not a universal solution due to the tendency of Newton's method to be sensitive to initial conditions.

Recall that Newton's method for a scalar equation  $f(x) = 0$  is based on linearization: if  $x$  is a current guess for the root, we try to find an updated guess  $x + \delta x$  by Taylor expansion:

$$f(x + \delta x) \approx f(x) + f'(x)\delta x. \quad (10.48)$$

Neglecting the higher order terms gives a linear equation for  $\delta x$ ,

$$\delta x = -f'(x)^{-1} f(x). \quad (10.49)$$

This gives an iteration  $x^{(n)} \rightarrow x^{(n+1)}$ .

The equation  $F_{ai}(\phi) = 0$  is a vector equation, with vector index  $\alpha = (ia)$ . We write our iteration as

$$\phi^{(n+1)} = \phi^{(n)} e^{\eta - \eta^H}. \quad (10.50)$$

That is,

$$\phi_p^{(n+1)} = \sum_q \phi_q^{(n)} (e^{\eta - \eta^H})_{qp}. \quad (10.51)$$

We need to find an equation for  $\eta$ , which is the unknown akin to  $\delta x$  above, in each iteration.

Since now we have  $\phi$  as a function of  $\eta$ , we write  $F_{ia}(\eta)$ , which we Taylor expand to first order around  $\eta = 0$ . Since  $\eta$  is complex, we will obtain terms that depend on both  $\eta_{bj}$  and  $\eta_{bj}^*$ . The Taylor expansion to first order is obtained by expanding all the bra and ket SPFs to first order,

$$\delta |\phi_i\rangle = \sum_a \eta_{ai} |\phi_a\rangle, \quad \delta |\phi_a\rangle = - \sum_i \eta_{ai}^* |\phi_i\rangle. \quad (10.52)$$

$$\delta \langle \phi_i| = \sum_a \eta_{ai}^* \langle \phi_a|, \quad \delta \langle \phi_a| = - \sum_i \eta_{ai} \langle \phi_i|. \quad (10.53)$$

$$F_{ai}(\eta) \approx F_{ai}(0) + \sum_{bj} \eta_{bj} K_{bj,ai} + \sum_{bj} \eta_{bj}^* L_{bj,ai}. \quad (10.54)$$

Here,

$$K_{bj,ai} = -\delta_{ba} \left( \sum_k \rho_i^k h_k^j + \sum_{klm} \rho_{im}^{kl} w_{kl}^{jm} \right) + \rho_i^j h_b^a + \sum_{kl} \rho_{ik}^{jl} (w_{bl}^{ak} - w_{lb}^{ak}) \quad (10.55a)$$

$$L_{bj,ai} = \sum_{kl} \rho_{ij}^{kl} w_{kl}^{ab} \quad (10.55b)$$

See Exercise 10.6. These matrices are in general complex. The equation (10.54) may now be split into its real and imaginary parts, giving linear equations for  $\text{Re } \eta$  and  $\text{Im } \eta$  that may be solved with linear algebra routines. We can then compute the next iteration of  $\phi$  using Eq. (10.50).

## 10.2.7 Complete active space self-consistent field (CAS-SCF)

Our treatment of MCHF so far has been generic. Specializing to electronic systems and molecular systems, it is common to let the SPFs be spin-orbitals of restricted type,

$$\phi_{p,\sigma}(\vec{r}, s) = \varphi_p(\vec{r}) \chi_\sigma(s).$$

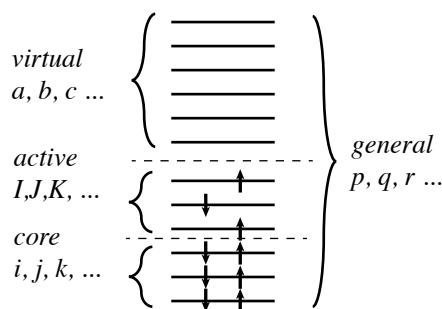


Figure 10.2: Illustration of the CAS partitioning of spin-orbitals. Core orbitals are always doubly occupied, while active/valence orbitals can have any occupation. Virtual orbitals have zero occupations.

A Slater determinant  $|\Phi_\mu\rangle$  has orbitals either empty, singly occupied, or doubly occupied. Our specification so far has that there are no constraints on the occupancy.

In the co-called complete active space (CAS) specification, we group the orbitals  $\varphi_p$  into *core* orbitals  $\varphi_i$  that are always doubly occupied, and *valence/active* orbitals  $\varphi_I$ , that can have any occupancy (zero, one, or two), and finally virtual orbitals  $\varphi_a$ , which are always unoccupied. See Figure ?? . The selection of a core orbital space greatly reduces the dimension of the amplitude expansion coefficients  $A$ , while still being still flexible enough to treat very complex molecules.

When each  $\phi_i$  is expanded in an atomic basis, we obtain the *complete active space self-consistent field method* (CAS-SCF). Another benefit of the CAS space is that it yields approximate eigenfunctions that are eigenfunctions also of total electron spin, while also treating other forms of static correlation.

It is also possible to do an unrestricted version of CAS-SCF, where the orbitals are different for different spins. This does *not* give spin eigenfunctions, but we can get better energies. To see this, compare RHF with UHF, which is a special case of CAS-SCF. In the trivial case of  $N/2$  core orbitals and zero active/virtual orbitals, we have only one possible determinant. The CAS wavefunction collapses to  $|\Psi_{\text{CAS}}\rangle = A_0 |\Phi_0\rangle$ , a single determinant of  $N/2$  doubly occupied orbitals. This is the restricted Hartree–Fock wavefunction. The derivation for UHF is similar.

In practical situations, CAS-SCF is often used as an “orbital generator” for more complex methods, just like Hartree–Fock. One then has a small CAS space, and include the found virtual orbitals in, say, a CI calculation.

## 10.2.8 Exercises

**Exercise 10.1.** Show that

$$|\Phi_\mu\rangle = e^{\hat{S}} |\hat{\Phi}_\mu\rangle, \quad (10.56)$$

see the text for definitions. △

**Exercise 10.2.** Let  $|\Psi\rangle$  be a wavefunction, and let  $\rho_p^q$  and  $\rho_{pr}^{qs}$  be the corresponding reduced one- and twobody density matrices. Show that the expectation value of the Hamiltonian  $\hat{H} = \hat{H}_0 + \hat{W}$  becomes

$$E = \sum_{qp} \rho_p^q h_q^p + \frac{1}{2} \sum_{prqs} \rho_{pr}^{qs} w_{qs}^{pr}. \quad (10.57)$$



In the case of the MCHF wavefunction, show that  $h_p^q$  is a function of the matrix  $S$  and not of the amplitude vector  $A$ .  $\triangle$

**Exercise 10.3.** Let  $|\Phi\rangle$  be the Hartree–Fock determinant viewed as a MCHF wavefunction. Explain that the number of active SPFs is the number of electrons, and that there is only one Slater determinant in the corresponding many-electron space. Compute the reduced density matrices  $\rho_i^j$  and  $\rho_{ik}^{jl}$ , and show that

$$\rho_i^j = \delta_{ji}, \quad \rho_{ik}^{jl} = \delta_{ji}\delta_{lk} - \delta_{jk}\delta_{li}. \quad (10.58)$$

Compute the total MCHF energy, and show that it coincides with the HF energy.

Show that the MCHF equations reduce to

$$\Omega = \hat{Q}\hat{f}|\phi_i\rangle.$$

Show that  $\Omega = 0$  if and only if there are constants  $\lambda_{ji}$  such that

$$\hat{f}|\phi_i\rangle = \sum_j \lambda_{ji} |\phi_j\rangle,$$

i.e., the non-canonical HF equations. (Hint: use  $\hat{Q}|\psi\rangle = 0$  implies  $|\psi\rangle = \hat{P}|\psi\rangle$ .)  $\triangle$

**Exercise 10.4.** In this exercise, we determine the variational equations of the MCHF energy with respect to the amplitudes.

a) Prove that

$$\frac{\partial E_{\text{MCHF}}}{\partial A_\mu^*} = \frac{1}{A^H A} [(KA)_\mu - E_{\text{MCHF}} A_\mu], \quad (10.59)$$

where  $K_{\mu\nu} = \langle \Phi_\mu | \hat{H} | \Phi_\nu \rangle$ .

b) Show that

$$\frac{\partial E_{\text{MCHF}}}{\partial A_\mu} = \frac{1}{A^H A} [(A^H K)_\mu - E_{\text{MCHF}} A_\mu^H]. \quad (10.60)$$

c) Show that the two equations are complex conjugates of each other if and only if the Hamiltonian  $\hat{H}$  is Hermitian.  $\triangle$

**Exercise 10.5.** Show that if  $S$  is a matrix on block form

$$S = \begin{pmatrix} S^{\text{act}} & 0 \\ 0 & S^{\text{vir}} \end{pmatrix},$$

then  $e^S$  is on the form

$$e^S = \begin{pmatrix} e^{S^{\text{act}}} & 0 \\ 0 & e^{S^{\text{vir}}} \end{pmatrix}.$$

$\triangle$

**Exercise 10.6.** Derive the MCHF Newton iteration matrices (10.55).  $\triangle$

## **Chapter 11**

# **Introduction to time-dependent theory (Not lectured 2017)**

This chapter is not written yet.

This chapter will contain methods for solving the time-dependent Schrödinger equation based on the wavefunction parameterizations we have encountered so far. The equations of motion are obtained via time-dependent variational principles.

We also introduce linearized equations of motion for excited states, for example TDHF or EOMCC.

## **Appendix A**

# **Mathematical supplement**

## A.1 Calculus of variations

### A.1.1 Functionals

In the calculus of variations, we compute the extrema of a possibly nonlinear *function of a function*. Such objects are often called *functionals*. Thus, a functional  $F[u]$  takes some *function*  $u$  and produces a *number*. One can think of  $F$  depending on infinitude of function values  $u(x)$ . In the case of the energy expectation value, the  $N$ -body wavefunction  $|\Psi\rangle$  is mapped to the number

$$\mathcal{E}[|\Psi\rangle] = \langle\Psi|\hat{H}|\Psi\rangle / \langle\Psi|\Psi\rangle.$$

Suppose we expand the wavefunction in a basis, say, a Slater determinant basis,

$$|\Psi\rangle = \sum_I A_I |\Phi_I\rangle.$$

Then,  $\mathcal{E}$  becomes a function of the vector  $\vec{A}$ , a possibly infinite set of coefficients. This may be an easier way to think of a functional: a function that depends on  $K$  variables, where  $K$  may be infinite.

A functional can also depend on more than one function. In Hartree–Fock theory, the energy functional depends on  $N$  single-particle functions  $\phi_i, i = 1, \dots, N$ . Moreover, the Hartree–Fock Lagrangian function that we *actually* optimize is a functional that also depends on a matrix  $\lambda = [\lambda_{ij}]$  of Lagrange multipliers,  $\mathcal{L} = \mathcal{L}[\phi_1, \dots, \phi_N, \lambda]$ . Given expansions of the  $\phi_i$  as  $\phi_i(x) = \sum_p \chi_p(x) U_{ip}$ , we see that  $\mathcal{L}$  becomes a function of the matrix  $U$  and the matrix  $\lambda$ . Thus, functionals are not too different from ordinary functions of a vectors.

How do we go about computing the extrema of a functional? A function of a single real variable has an intuitive notion of a local extremum, and most readers probably have an intuitive notion of extrema of two-variable functions as well. But if we go to higher dimensions (or infinite dimensions!) it becomes more complicated.

We will therefore introduce the concept of a *directional derivative* in a rather informal way. This is very handy, and allows us to read off the condition for an extremum in a straight-forward manner. This framework is called *the calculus of variations*, since we are computing the “variation in  $F[u]$ ” with respect to arbitrary “variations  $\delta u$  of the function  $u$ ”.

### A.1.2 Functions of one real variable

Consider first a simple function  $F : I \rightarrow \mathbb{R}$ ,  $I \subset \mathbb{R}$  being an interval. Suppose  $x_0 \in I$ . Assuming that  $F$  can be differentiated at least twice, we can compute a second-order Taylor expansion around  $x_0$ , viz,

$$F(x_0 + \epsilon) \approx F(x_0) + \epsilon F'(x_0) + \frac{1}{2} \epsilon^2 F''(x_0). \quad (\text{A.1})$$

The error in this approximation vanishes as  $\epsilon \rightarrow 0$ .

The condition for an extremum at  $x_0$  is  $F'(x_0) = 0$ . The second-order term tells us the nature of the extremum: if  $F''(x_0) > 0$  then  $x_0$  is a local minimum. If  $F''(x_0) < 0$  then  $x_0$  is a local maximum. Finally, if  $F''(x_0) = 0$ , we cannot determine right away if we have a maximum or minimum. We may have neither, as for  $F(x) = x^3$ , where  $x_0 = 0$  is a saddle point. A minimum and a saddle point is illustrated in Fig. A.1.

### A.1.3 Functions of two real variables

Consider yourself in a landscape of mountains and valleys. The elevation is  $F(x, y)$ . You are trying to find, say, a local minimum  $(x_0, y_0)$  of elevation. On a map, a local minimum will show up as successively

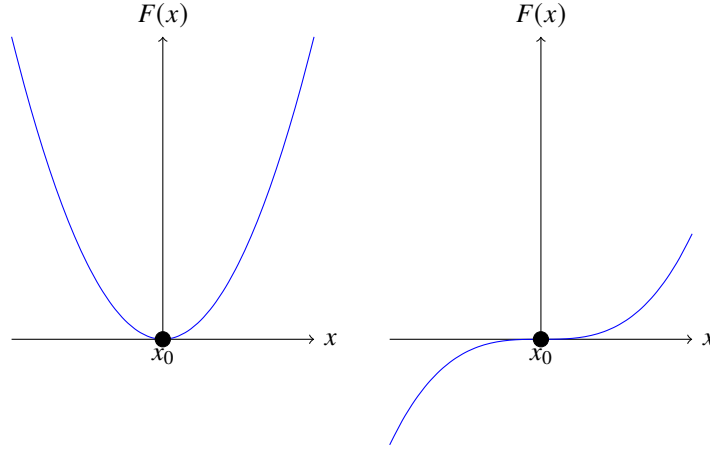


Figure A.1: Simple functions of one real variables with a local minimum ( $F''(x_0) > 0$ ) (left) and a saddle point ( $F''(x_0) = 0$ ) (right)..

smaller closed curves of equal elevation, see Fig. A.2. (The same is true for a maximum, and a saddle point is a crossing of lines of equal elevation.) We now observe, that if you move in a direction  $\eta = (\delta x, \delta y) \neq 0$  from the local minium, you *will always walk uphill*, that is, the function

$$f(\epsilon) = F(x_0 + \epsilon \delta x, y_0 + \epsilon \delta y)$$

has a local minimum at  $\epsilon = 0$ , irrespective of  $\eta$ . If you were standing on a mountaintop (a local maximum) you would always walk downhill, and  $f(\epsilon)$  would always have a local maximum at  $\epsilon = 0$ .

Finally, if you are standing between two mountaintops to the east and west, and looking down at valleys to the south and north, you are standing on a saddle point. You are walking downhill if you go north or south, but uphill if you go east or west:  $f(\epsilon)$  has a local minimum for some  $\eta$ , and a maximum for other  $\eta$ .

We see that, at least intuitively, we can determine wheter  $F$  has a local extremum at  $(x_0, y_0)$  by studying the behaviour of  $f(\epsilon)$ , for all possible choices of  $\eta$ . We now prove this claim:

Let us compute the Taylor expansion of  $f(\epsilon)$ :

$$\begin{aligned} f(\epsilon) &\approx f(0) + \epsilon f'(0) + \frac{1}{2} \epsilon^2 f''(0) \\ &= F(x_0, y_0) + \epsilon \nabla F(x_0, y_0)^T \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} + \frac{1}{2} \epsilon^2 (\delta x \ \delta y) H(x_0, y_0) \begin{pmatrix} \delta x \\ \delta y \end{pmatrix}. \end{aligned} \quad (\text{A.2})$$

We used the chain rule, and introduced the gradient and the Hessian matrix  $H$ , given by

$$\nabla F(x_0, y_0) = \begin{pmatrix} \frac{\partial F(x_0, y_0)}{\partial x} \\ \frac{\partial F(x_0, y_0)}{\partial y} \end{pmatrix} \quad (\text{A.3})$$

and

$$H(x_0, y_0) = \begin{pmatrix} \frac{\partial^2 F(x_0, y_0)}{\partial x^2} & \frac{\partial^2 F(x_0, y_0)}{\partial x \partial y} \\ \frac{\partial^2 F(x_0, y_0)}{\partial y \partial x} & \frac{\partial^2 F(x_0, y_0)}{\partial y^2} \end{pmatrix}. \quad (\text{A.4})$$

Now,  $F$  has an extremum at  $(x_0, y_0)$  if and only if  $\nabla F(x_0, y_0) = 0$ , while  $f(\epsilon)$  has an extremum at  $\epsilon = 0$  if and only if the second term in Eq. (A.2) vanishes. But if  $\nabla F(x_0, y_0)^T \eta = 0$  for all  $\eta \neq 0$ , then clearly  $\nabla F(x_0, y_0) = 0$  and vice versa. QED.

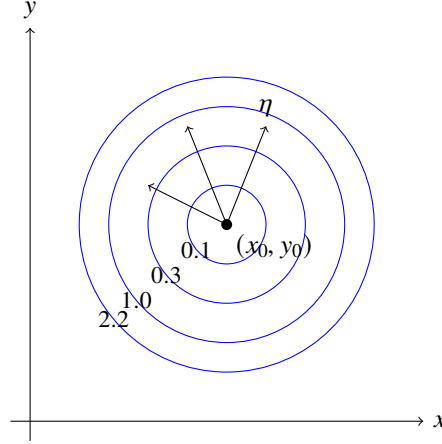


Figure A.2: The condition for a local minimum  $(x_0, y_0)$  for a function  $F(x, y)$ : in all directions  $\eta \neq 0$  you walk uphill from  $(x_0, y_0)$ .

We introduce the *directional derivative* of  $F$  at  $(x_0, y_0)$  in the direction  $\eta = (\delta x, \delta y)$ ,

$$F'(x_0, y_0; \eta) \equiv \left. \frac{d}{d\epsilon} F(x_0 + \epsilon \delta x + y_0 + \epsilon \delta y) \right|_{\epsilon=0} \quad (\text{A.5})$$

which is precisely the second term in Eq. (A.2),

$$f(\epsilon) \approx F(x_0, y_0) + \epsilon F'(x_0, y_0; \eta) + \frac{1}{2} \epsilon^2 \eta^T H(x_0, y_0) \eta. \quad (\text{A.6})$$

Thus, the extremum condition is equivalent to  $F'(x_0, y_0; \eta) = 0$  for all  $\eta \neq 0$ .

What about the nature of the extremum? If

$$\eta^T H(x_0, y_0) \eta > 0 \quad (\text{A.7})$$

for all possible directions  $\eta$ , we have a local minimum. This is precisely the condition that  $H(x_0, y_0)$  is a positive definite matrix. Since  $H(x_0, y_0)$  is a symmetric matrix, this is equivalent to all the eigenvalues being positive. Thus,  $f(\epsilon)$  must have a local minimum at  $\epsilon = 0$  for every  $\eta \neq 0$ .

Similarly, if  $H(x_0, y_0)$  is negative definite,

$$\eta^T H(x_0, y_0) \eta < 0, \quad \forall \eta \quad (\text{A.8})$$

then we have a local maximum. However, if  $H(x_0, y_0)$  is neither positive nor negative definite, we cannot say whether we have a maximum or minimum. We may in fact have a saddle point, as in the case of standing between mountains and valleys.

#### A.1.4 Extremalization of a functional

The concept of the directional derivative is of course valid for more than two dimensions. For a function  $F : \mathbb{R}^n \rightarrow \mathbb{R}$ , the localization of an extremum can be formulated as: find  $x_0 \in \mathbb{R}^n$  such that the directional derivative vanishes for every nonzero  $\eta \in \mathbb{R}^n$ :

$$F'(x_0; \eta) = \left. \frac{d}{d\epsilon} F(x_0 + \epsilon \eta) \right|_{\epsilon=0} = 0, \quad \forall \eta \in \mathbb{R}^n, \eta \neq 0. \quad (\text{A.9})$$

This condition is equivalent to  $\nabla F(x_0)^T = 0$ .

Turning to a *functional*  $F[u]$  for some function  $u$ , or set of functions, the directional derivative in the direction of the *function*  $\eta$  is in principle straightforward:

$$F'[u; \eta] = \left. \frac{d}{d\epsilon} F[u + \epsilon\eta] \right|_{\epsilon=0}. \quad (\text{A.10})$$

Computing  $F[u + \epsilon\eta]$  as a series in  $\epsilon$  is usually straightforward, allowing an expression for  $F'[u; \eta]$  to be read off. Typically, this leads to a *differential equation*: the variational principle gave us the Schrödinger equation, while extremalization of the Hartree–Fock energy gave us the Hartree–Fock equations.

The term “calculus of variations” is historical, and comes from the idea that we are “computing infinitesimal variations  $\delta F[u]$  in the functional under infinitesimal variations  $\delta u$  of the function” in all possible ways, i.e., a different way of saying that we are computing directional derivatives.

# Bibliography

- [1] S. Raimes. *Many-electron theory*. North-Holland, Amsterdam, Netherlands, 1972.
- [2] E.K.U. Gross, E. Runge, and O. Heinonen. *Many-Particle Theory*. Adam Hilger, Bristol, Philadelphia and New York, 1991.
- [3] A. Szabo and N.S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Dover, New York, USA, 1996.
- [4] J.M. Leinaas and J. Myrheim. On the theory of identical particles. *Il Nuovo Cimento*, 37:1, 1977.
- [5] F.E. Harris, H.J. Monkhorst, and D.L. Freeman. *Algebraic and Diagrammatic Methods in Many-Fermion Theory*. Oxford, New Work, 1992.
- [6] R. Helgaker, P. Jørgensen, and Olsen. *J. Molecular Electronic-Structure Theory*. Wiley, Chichester, UK, 2002.
- [7] J. Paldus and Čížek. Time-independent diagrammatic approach to perturbation theory of fermion systems. *Adv. Quant. Chem.*, 9:105, 1975.
- [8] G.F. Giuliani and G. Vignale. *Quantum Theory of the Electron Liquid*. Cambridge, Cambridge, UK, 2005.
- [9] M. Moshinsky and Y.F. Smirnov. *The Harmonic Oscillator in Modern Physics*. Harwood, Amsterdam, Netherlands, 1996.
- [10] T. D. Crawford and H.F. Schaefer III. An introduction to coupled cluster theory for computational chemists. *Rev. Comp. Chem.*, 14:33, 2000.
- [11] I. Shavitt and R. J. Bartlett. *Many-body methods in chemistry and physics: MBPT and Coupled-Cluster Theory*. Cambridge, 2009.