

Wavefunction-based Correlation Methods:

Overview of Mathematical Tools & Second Quantization

Zhendong Li

College of Chemistry
Beijing Normal University

April 8, 2021

Contents

- 1 Introduction
- 2 Mathematical tools
 - Linear algebra
 - Numerical optimization
 - Nonlinear equation
- 3 Second quantization

Introduction

This lecture will give an overview of electron correlation methods (beyond Hartree-Fock/mean-field theory) for solving the electronic Schrödinger equation

$$\hat{H}_{\text{el}}\Psi(x_1, \dots, x_N) = \Psi(x_1, \dots, x_N)E, \quad (1)$$

where \hat{H}_{el} within the Bohn-Oppenheimer approximation reads

$$\hat{H}_{\text{el}} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{ne}, \quad (2)$$

$$\hat{T}_e = \sum_i -\frac{1}{2}\nabla_i^2, \quad (3)$$

$$\hat{V}_{ee} = \sum_{i>j} \frac{1}{r_{ij}}, \quad (4)$$

$$\hat{V}_{ne} = \sum_i \sum_A -\frac{Z_A}{r_{iA}}. \quad (5)$$

The nuclear-electron attraction \hat{V}_{ne} defines **the specific molecule**.

Introduction

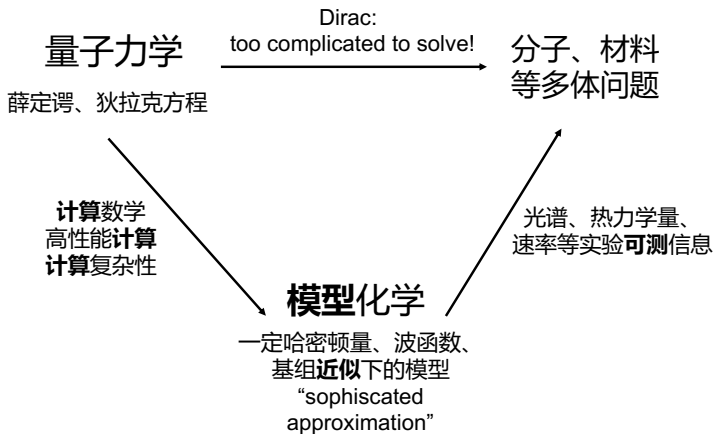
The challenging part is the electron-electron interaction \hat{V}_{ee} , which prevents the solution by separation of variables.

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much **too complicated to be soluble**. It therefore becomes desirable that **approximate practical methods** of applying quantum mechanics should be developed, which can lead to an explanation of the main features of complex atomic systems without too much computation.^a

^aPaul Adrien Maurice Dirac. "Quantum mechanics of many-electron systems". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733.

The field of electron correlation is to design approximate methods that are both **accurate** and **efficient**!

Introduction



Introduction

Nobel Prize in Chemistry (related with electronic structure): <https://www.nobelprize.org/prizes/lists/all-nobel-prizes-in-chemistry/>

- **1966 Robert S. Mulliken**

"for his fundamental work concerning chemical bonds and the electronic structure of molecules by the molecular orbital method"

- **1981 Kenichi Fukui & Ronald Hoffman**

"for their theories, developed independently, concerning the course of chemical reactions"

- **1998 Walter Kohn & John A. Pople**

"for his development of the density functional theory"

"for his development of computational methods in quantum chemistry"

- **2013 Martin Kaplus & Micheal Levitt & Arie Warshel**

"for the development of multiscale models for complex chemical systems"

More recent theoretical achievements: International Academy of Quantum Molecular Science (IAQMS) <https://www.iaqms.org/awards.php>

Introduction

Classification of electron correlation methods by fundamental principles:

Ground states

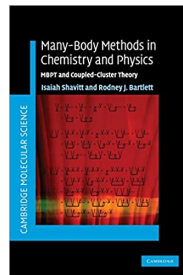
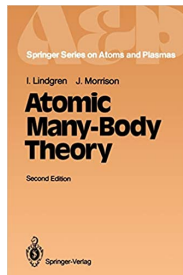
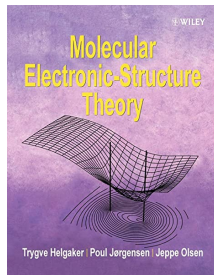
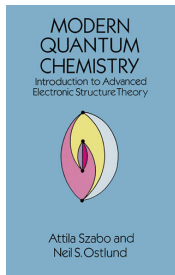
- 1 Variational principle: HF, CASSCF, SR/MR-CI, SCI, DMRG, VMC
- 2 Perturbation theory: MBPT (MP2, MRPT, etc.), GF methods (RPA)
- 3 Similarity transformation: CC
- 4 Imaginary-time formulation: Projector MC (DMC, FCIQMC)

Excited states

- 1 Variational principle: CIS, SA-CASSCF, MRCI, orthogonality constraint
- 2 Equation of motion: RPA, EOM-CC
- 3 Linear response theory: TDHF, BSE

We will only discuss the basic ideas behind those highlighted methods.

Introduction



Introduction

Recall the variational principle:

$$\langle \hat{H} \rangle \triangleq \frac{\langle \Psi | \hat{H} | \Psi \rangle}{\langle \Psi | \Psi \rangle} \geq E_0. \quad (6)$$

Examples

- Heitler-London (1927): valence-bond (VB) trial wavefunction:

$$\Psi_{H_2}(x_1, x_2) = \mathcal{N}[\psi_{H_A}(r_1)\psi_{H_B}(r_2) + \psi_{H_A}(r_2)\psi_{H_B}(r_1)]\Theta_{0,0}, \quad (7)$$

$$\Theta_{0,0} = \frac{1}{\sqrt{2}}(\alpha(1)\beta(2) + \beta(1)\alpha(2)). \quad (8)$$

The calculation explains the formation of chemical bond, which cannot be explained in the 19's century without the advent of quantum mechanics.

- D. R. Hartree (1927): product wavefunction for atomic electronic structure, later becomes Hartree-Fock taking into account of **antisymmetry principle**.

Introduction

Proof

$|\Psi\rangle = \sum_{i=0} c_i |E_i\rangle$ with $E_0 \leq E_1 \leq \dots$, then

$\langle \hat{H} \rangle = \frac{\sum_{i=0} |c_i|^2 E_i}{\sum_{i=0} |c_i|^2} \geq \frac{\sum_{i=0} |c_i|^2 E_0}{\sum_{i=0} |c_i|^2} = E_0$. The equality holds if and only if $c_0 = 1$, i.e., $|\Psi\rangle = |E_0\rangle$.

Examples

- Linear variational theory: $|\Psi_{CI}\rangle = \sum_I c_I |\Phi_I\rangle$, where the set of basis functions $\{|\Phi_I\rangle\}$ is fixed.
- Nonlinear variational theory: Hartree-Fock, CASSCF, etc.

Question: How to realize theories in practice? We need **numerical methods**.

Flowchart: Problem \rightarrow Theory \rightarrow Equations \rightarrow Implementation \rightarrow Applications

Contents

1 Introduction

2 Mathematical tools

- Linear algebra
- Numerical optimization
- Nonlinear equation

3 Second quantization

Introduction

We will cover three topics very briefly:

- **Linear algebra** (used in HF and CI)
- **Numerical optimization** (used in CASSCF)
- **Nonlinear equation** (used in CC)

They are essential in electron correlation methods for solving stationary Schrödinger equation. Other important topics, which are less common, include

- **Monte-Carlo methods** (used in QMC)
- **Ordinary differential equation** (used in time-dependent problems)

Some good references:

- *Numerical Python: A Practical Techniques Approach for Industry*
<http://jrjohansson.github.io/numericalpython.html>
- *Numerical Recipes*: <http://numerical.recipes/book/book.html>

Eigenvalue problems

(Generalized) eigenvalue problem is defined by

$$\mathbf{H}\mathbf{c} = \mathbf{S}\mathbf{c}E. \quad (9)$$

It appears in many places in quantum chemistry:

- Hermitian case:
 - Hartree-Fock-Roothaan equation: $\mathbf{F}\mathbf{C} = \mathbf{S}\mathbf{C}E$
 - Configuration interaction (CI): CIS, CISD, etc. ($\mathbf{S} = \mathbf{I}$)
- Non-Hermitian case:
 - TD-DFT, Bethe-Salpeter equation (BSE), EOM-CC

Eigenvalue problems

Hermitian case is more well-behaved (assuming $\mathbf{S} = \mathbf{I}$) in the sense that:

- Eigenvalues are real;
- Eigenvectors can be made orthonormal: \mathbf{C} is unitary ($\mathbf{C}^\dagger \mathbf{C} = \mathbf{I}$).

Non-Hermitian matrices (appeared in TD-DFT) can have non-real eigenvalues.

$$\begin{bmatrix} \Delta + K & K \\ -K & -(\Delta + K) \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = \begin{bmatrix} X \\ Y \end{bmatrix} \omega \quad (10)$$

Eigenvalue:

$$\det \begin{bmatrix} \Delta + K - \omega & K \\ -K & -(\Delta + K) - \omega \end{bmatrix} = 0 \Rightarrow \omega = \pm \sqrt{\Delta(\Delta + 2K)} \quad (11)$$

Condition for real eigenvalues: $\Delta(\Delta + 2K) \geq 0$.

Eigenvalue problems

Example: Pauli matrices

$$\sigma_x = X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad \sigma_y = Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \quad \sigma_z = Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad (12)$$

Eigenvectors:

$$\mathbf{U}_x = H = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}, \quad \mathbf{U}_y = SH = \begin{bmatrix} 1 & 0 \\ 0 & i \end{bmatrix} H, \quad \mathbf{U}_z = I, \quad (13)$$

such that

$$HXH = Z, \quad HS^\dagger YSH = Z. \quad (14)$$

The matrices H and S are referred as Hadamard and phase gates, respectively, in quantum computing literature.

Eigenvalue problems

Given $N \times N$ matrix \mathbf{H} , the computational cost of finding its eigenvalue scales as

- Full diagonalization: $O(N^3)$ [e.g., `dsyev` in LAPACK, `eigh` in numpy] - usually used in diagonalization of the Fock matrix in SCF.
- Iterative methods (requiring $\sigma = \mathbf{H}\mathbf{x}$): $O(N_{iter}N^2)$ [e.g., Power method, Lanczos algorithm, Davidson algorithm¹] - widely used in finding approximate eigenstates in CI and TD-DFT, because usually only a few low-lying states are of interest.

¹Ernest R Davidson. "The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices". In: *Journal of Computational Physics* 17 (1975), pp. 87–94.

Matrix perturbation theory

Perturbation theory is another useful technique to obtain approximate solutions or perform analysis. Later, we will use it in the context of many-body problems to develop MP2 and MRPT2.

It can be classified as follows:

- Time-dependent perturbation theory
- Time-independent perturbation theory
 - nondegenerate PT
 - (quasi-)degenerate PT

Here, we will focus on the nondegenerate PT for simplicity. We assume a partition $H = H_0 + V$, where V represents external perturbations (electric, magnetic field, SOC, etc.).

Matrix perturbation theory

In the simplest case, nondegenerate PT can be considered as the **Taylor expansion** for eigenvalues and eigenvectors. Consider the following problem for the two level system (TLS \equiv **qubit**)

$$H_0 = \begin{pmatrix} -\frac{\Delta}{2} & 0 \\ 0 & \frac{\Delta}{2} \end{pmatrix} = -\frac{\Delta}{2} \mathbf{Z}, \quad (15)$$

$$V = \begin{pmatrix} 0 & v \\ v & 0 \end{pmatrix} = v \mathbf{X}. \quad (16)$$

We assume $\Delta, v > 0$ for convenience. In this case, the exact eigenvalue for $H = H_0 + V$ can be worked out explicitly (an exercise) as

$$E(v) = \pm \frac{1}{2} \sqrt{\Delta^2 + 4v^2} \quad (17)$$

Matrix perturbation theory

The Taylor expansion of the lowest eigenvalue $E_0(v) = -\frac{1}{2}\sqrt{\Delta^2 + 4v^2}$ leads to

$$E_0(v) = -\frac{\Delta}{2} - \frac{v^2}{\Delta} + \frac{v^4}{\Delta^3} + O(v^6). \quad (18)$$

The expansion converge if $4v^2 < \Delta^2$.

We now derive the second-order term by PT. The zeroth-order problem is simple to solve (**essential for PT!**),

$$|0^{(0)}\rangle = |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad E_0^{(0)} = \langle 0^{(0)} | H_0 | 0^{(0)} \rangle = -\frac{\Delta}{2}, \quad (19)$$

$$|1^{(0)}\rangle = |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad E_1^{(0)} = \langle 1^{(0)} | H_0 | 1^{(0)} \rangle = \frac{\Delta}{2}. \quad (20)$$

Matrix perturbation theory

To find $E_0^{(2)}$, we perform a similar expansion for the eigenvalue problem

$$(H_0 + V)|0(v)\rangle = |0(v)\rangle E_0(v) \quad (21)$$

to second order

$$H_0|0^{(0)}\rangle = |0^{(0)}\rangle E_0^{(0)} \quad (22)$$

$$H_0|0^{(1)}\rangle + V|0^{(0)}\rangle = |0^{(1)}\rangle E_0^{(0)} + |0^{(0)}\rangle E_0^{(1)} \quad (23)$$

$$H_0|0^{(2)}\rangle + V|0^{(1)}\rangle = |0^{(2)}\rangle E_0^{(0)} + |0^{(1)}\rangle E_0^{(1)} + |0^{(0)}\rangle E_0^{(2)} \quad (24)$$

which can be rearranged into

$$(H_0 - E_0^{(0)})|0^{(1)}\rangle = -V|0^{(0)}\rangle + |0^{(0)}\rangle E_0^{(1)} \quad (25)$$

$$(H_0 - E_0^{(0)})|0^{(2)}\rangle = -V|0^{(1)}\rangle + |0^{(1)}\rangle E_0^{(1)} + |0^{(0)}\rangle E_0^{(2)} \quad (26)$$

Matrix perturbation theory

$$(H_0 - E_0^{(0)})|0^{(1)}\rangle = -V|0^{(0)}\rangle + |0^{(0)}\rangle E_0^{(1)} \quad (27)$$

$$(H_0 - E_0^{(0)})|0^{(2)}\rangle = -V|0^{(1)}\rangle + |0^{(1)}\rangle E_0^{(1)} + |0^{(0)}\rangle E_0^{(2)} \quad (28)$$

Applying $\langle 0|$ on both sides,

$$0 = -\langle 0|V|0\rangle + E_0^{(1)} \quad (29)$$

$$0 = -\langle 0|V|0^{(1)}\rangle + \langle 0|0^{(1)}\rangle E_0^{(1)} + E_0^{(2)} \quad (30)$$

Applying $\langle m|$ ($m > 0$) on both sides of the first-order equation,

$$(E_m^{(0)} - E_0^{(0)})\langle m|0^{(1)}\rangle = -\langle m|V|0^{(0)}\rangle \quad (31)$$

such that

$$\boxed{\langle m|0^{(1)}\rangle = \frac{\langle m|V|0\rangle}{(E_0^{(0)} - E_m^{(0)})}, \quad m > 0} \quad (32)$$

Matrix perturbation theory

Consequently,

$$\begin{aligned} E_0^{(2)} &= \langle 0|V|0^{(1)}\rangle - \langle 0|0^{(1)}\rangle E_0^{(1)} \\ &= \sum_{m \geq 0} \langle 0|V|m\rangle \langle m|0^{(1)}\rangle - \langle 0|0^{(1)}\rangle E_0^{(1)} \\ &= \sum_{m > 0} \langle 0|V|m\rangle \langle m|0^{(1)}\rangle \end{aligned} \tag{33}$$

which is independent of $\langle 0|0^{(1)}\rangle$. In sum,

$$\boxed{E_0^{(1)} = \langle 0|V|0\rangle} \tag{34}$$

$$\boxed{E_0^{(2)} = \sum_{m > 0} \frac{|\langle m|V|0\rangle|^2}{(E_0^{(0)} - E_m^{(0)})}} \tag{35}$$

These are important results, which we will use again in MP2.

Matrix perturbation theory

Remark: suppose $|\psi\rangle$ is a solution, then $|\varphi\rangle = e^{i\phi}|\psi\rangle$ (ϕ is real) is also a solution

$$\begin{aligned} |\varphi\rangle &= e^{i(\phi^{(1)} + \phi^{(2)} + \dots)} (|\psi^{(0)}\rangle + |\psi^{(1)}\rangle + |\psi^{(2)}\rangle + \dots) \\ &= (1 + i\phi^{(1)} + i\phi^{(2)} - \frac{1}{2}(\phi^{(1)})^2 + \dots) (|\psi^{(0)}\rangle + |\psi^{(1)}\rangle + |\psi^{(2)}\rangle + \dots) \\ &= |\psi^{(0)}\rangle \\ &+ |\psi^{(1)}\rangle + i\phi^{(1)}|\psi^{(0)}\rangle \\ &+ |\psi^{(2)}\rangle + i\phi^{(1)}|\psi^{(1)}\rangle + [i\phi^{(2)} - \frac{1}{2}(\phi^{(1)})^2]|\psi^{(0)}\rangle + \dots \\ &= |\varphi^{(0)}\rangle + |\varphi^{(1)}\rangle + |\varphi^{(2)}\rangle + \dots \end{aligned} \tag{36}$$

Therefore, the component $\langle 0|\varphi^{(1)}\rangle = \langle 0|\psi^{(1)}\rangle + i\phi^{(1)}$ is not uniquely determined, due to the arbitrariness of the global phase! The **intermediate** normalization refers to the choice $\langle 0|\varphi^{(n)}\rangle = 0$ ($n > 0$), which makes higher order PT simpler.

Matrix perturbation theory

Back to our two-level example,

$$E_0^{(1)} = \langle 0|V|0\rangle = 0 \quad (37)$$

$$E_0^{(2)} = \frac{|\langle 1|V|0\rangle|^2}{(E_0^{(0)} - E_1^{(0)})} = -\frac{v^2}{\Delta} \quad (38)$$

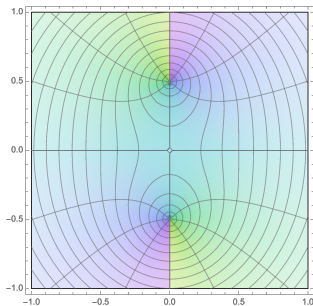
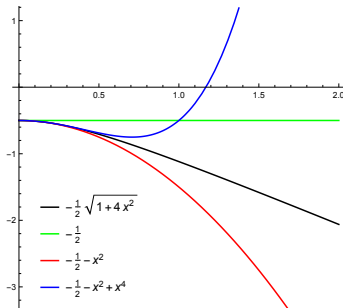
This agrees with the Taylor expansion

$$E_0(v) = -\frac{\Delta}{2} - \frac{v^2}{\Delta} + \frac{v^4}{\Delta^3} + O(v^6). \quad (39)$$

Matrix perturbation theory

Remarks: PT fails in the following cases (recall $E_0 = -\frac{1}{2}\Delta\sqrt{1 + (\frac{2v}{\Delta})^2}$)

- Large v : ComplexPlot (in Mathematica) for $-\frac{1}{2}\sqrt{1 + 4z^2}$ ($\Delta = 1$ and $v = z$)



- Small gap ($\Delta = 0$): degenerate PT

Numerical optimization

The lowest eigenvalue problem of a Hermitian matrix can be equivalently formulated as a minimization problem for the **Rayleigh quotient**,

$$R_H(x) = \frac{x^\dagger H x}{x^\dagger x} \quad (40)$$

which reaches its minimum value $R = \lambda_{\min}$ (smallest eigenvalue of H) when x is the corresponding eigenvector. More generally, the min-max theorem gives a variational characterization of eigenvalues of compact Hermitian operators on Hilbert spaces.

Theorem (Min-Max)

Let A be self-adjoint, and let $E_1 \leq E_2 \leq E_3 \leq \dots$ be the eigenvalues of A below the essential spectrum. Then

$$E_n = \min_{\psi_1, \dots, \psi_n} \max_{\psi} \{ \langle \psi, A\psi \rangle : \psi \in \text{span}(\psi_1, \dots, \psi_n), \|\psi\| = 1 \} \quad (41)$$

For details, see https://en.wikipedia.org/wiki/Min-max_theorem.

Numerical optimization

For simplicity, consider the real symmetric H , then the stationary condition of $R_H(x)$ is equivalent to the eigenvalue problem

$$\frac{\partial R_H(x)}{\partial x_k} = 2 \frac{1}{x^T x} [(Hx)_k - R_H(x)x_k] = 0. \quad (42)$$

In particular, for the lowest eigenstate c_0 , $R_H(x)$ achieves its minimum E_0 , as can be seen from the Hessian at the stationary point,

$$\left[\frac{\partial^2 R_H(x)}{\partial x_l \partial x_k} \right] \Big|_{x=c_0} = 2[H - E_0 I]_{lk} \succeq 0. \quad (43)$$

Therefore, to find E_0 and c_0 , we can use numerical optimization techniques to minimize $R_H(x)$ (**variational principle**). Note that $R_H(x) = R_H(\lambda x)$ is invariant with respect to the scale transformation, such that we can normalize the vector x during calculations.

Methods for numerical optimization

The methods for unconstrained optimization can be classified as

- Gradient free: Nelder-Mead
- Gradient-based: gradient descent (GD) method, conjugate gradient, ...
- Hessian-based: Newton's method
- Quasi-Newton methods: Broyden-Fletcher-Goldfarb-Shanno (BFGS)

which differ on the convergence properties, such as the rate of convergence, global/local convergence, etc.

Gradient descent

Gradient descent (GD): first proposed by Cauchy in 1847

- Feature: **first-order** (only use gradient) iterative optimization algorithm for finding a local minimum of a differentiable function
- Mathematical formulation:

$$\vec{x}^{(n+1)} = \vec{x}^{(n)} - \gamma \nabla f(\vec{x}^{(n)}), \quad (44)$$

where $\gamma > 0$ is the step size. For small enough γ , $f(\vec{x}^{(n+1)}) = f(\vec{x}^{(n)} - \gamma \nabla f(\vec{x}^{(n)})) \approx f(\vec{x}^{(n)}) - \gamma \nabla f(\vec{x}^{(n)}) \cdot \nabla f(\vec{x}^{(n)}) < f(\vec{x}^{(n)})$.

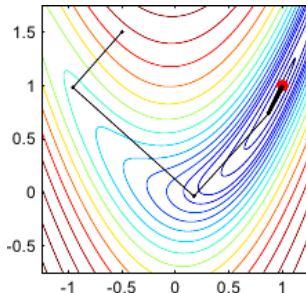
- Step size: γ can be found by (inexact) line search, e.g., $\min_{\gamma} f(\vec{x}^{(k)} + \gamma \vec{p}^{(k)})$ or set adaptively (referred as learning rate), e.g., see Nesterov's acceleration method.
- Extensions: the stochastic versions (SGD) are widely used in deep learning

Gradient descent

Rosenbrock function

$$f(x_1, x_2) = (1 - x_1)^2 + 100(x_2 - x_1^2)^2 \quad (45)$$

This pathological function has a narrow curved flat valley which contains the minimum. GD zigzags slowly with small step sizes towards the minimum.



Newton's Method

The central idea is to use second-order approximation around $x^{(n)}$

$$f(x^{(n)} + \delta) \approx f(x^{(n)}) + g_n^T \delta + \frac{1}{2} \delta^T H_n \delta \quad (46)$$

where $g_n = \nabla f(x^{(n)})$ and $H_n = [\nabla \nabla f(x^{(n)})]$ represent the gradient and Hessian, respectively. Minimizing the right hand side leads to a linear equation

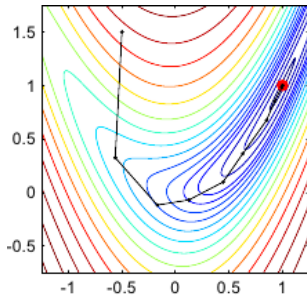
Newton's step

$$H \delta = -g_n. \quad (47)$$

The update is $x^{(n+1)} = x^{(n)} + \gamma \delta$. Near the local minima, the convergence is quadratic. Outside of the neighbourhoods of the local minima, there is no guarantee that Newton's method will converge. Besides, the Newton's method can also converge to saddle points.

Newton's Method

Performance for the Rosenbrock function:



Quasi-Newton algorithms such as BFGS use approximate Hessian to avoid the expensive evaluation of Hessian.

Numerical optimization

Message: in most cases, you can find a library implemented the optimization algorithm you want!

Reference:

- Brief introduction: <https://www.brnt.eu/phd/node10.html>
- Book by Nocedal and Wright:
<http://users.iems.northwestern.edu/~nocedal/book/num-opt.html>.
- Course on optimization: <https://www.stat.cmu.edu/~ryantibs/convexopt-S15/>
- `scipy.optimize.minimize`: <https://docs.scipy.org/doc/scipy/reference/generated/scipy.optimize.minimize.html>
- Test functions for optimization:
https://en.wikipedia.org/wiki/Test_functions_for_optimization
- Stochastic gradient descent:
https://en.wikipedia.org/wiki/Stochastic_gradient_descent

Nonlinear equation

Apart from eigenvalue and optimization problems, another numerical task frequently appeared in quantum chemistry is solving nonlinear equations, as we will see in coupled-cluster theory. A general nonlinear equation can be expressed as

$$\vec{f}(\vec{x}) = 0 \quad (48)$$

One useful idea to find the solution \vec{x} is **fixed point iteration**. In the simplest case, one can design a iterative formula $\vec{x}^{(k+1)} = \vec{f}(\vec{x}^{(k)}) + \vec{x}^{(k)}$. However, such naive scheme may not converge. Better scheme needs to use the information of \vec{f} . In the **Newton's method**, \vec{f} is approximate locally as

$$\vec{f}(\vec{x}) = \vec{f}(\vec{x}^{(k)} + [\vec{x} - \vec{x}^{(k)}]) \approx \vec{f}(\vec{x}^{(k)}) + Df(\vec{x}^{(k)})(\vec{x} - \vec{x}^{(k)}) = 0 \quad (49)$$

where $J_{ik} = [Df]_{ik} = \frac{\partial f_i}{\partial x_k}$ is the **Jacobian matrix**, leading to the following iterative formula

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} - [Df(\vec{x}^{(k)})]^{-1} \vec{f}(\vec{x}^{(k)}). \quad (50)$$

Contents

- 1 Introduction
- 2 Mathematical tools
 - Linear algebra
 - Numerical optimization
 - Nonlinear equation
- 3 Second quantization

Many-electron problem

Under the Born-Oppenheimer approximation (BOA), the time-independent nonrelativistic Schrödinger equation (SE) reads

$$\hat{H}_{\text{el}}\Psi = E\Psi \quad (51)$$

$$\hat{H}_{\text{el}} = \hat{H}_1 + \hat{H}_2 + E_{\text{nuc}}, \quad E_{\text{nuc}} = \sum_{A < B}^M \frac{Z_A Z_B}{R_{AB}} \quad (52)$$

$$\hat{H}_1 = \sum_{i=1}^N \hat{h}(i), \quad \hat{h}(i) = -\frac{1}{2}\nabla_i^2 + \sum_{A=1}^M \frac{Z_A}{r_{iA}} \quad (53)$$

$$\hat{H}_2 = \sum_{i < j}^N \frac{1}{r_{ij}} \quad (54)$$

This will be referred as the SE in **coordinate space** or **first quantization**.

N -electron Hilbert space

Given a set of orthonormal spin-orbitals $\{\phi_k(x)\}_{k=1}^K$,

$$\langle \phi_k | \phi_l \rangle = \delta_{kl}. \quad (55)$$

The N -electron Hilbert space \mathcal{H}_N can be expanded by the set of orthonormal Slater determinants

$$\{|\phi_{k_1}(x_1)\phi_{k_2}(x_2)\cdots\phi_{k_N}(x_N)\rangle : k_1 < k_2 < \cdots < k_N\} \quad (56)$$

The restriction $k_1 < k_2 < \cdots < k_N$ is due to the **antisymmetry principle**
 $|\phi_{k_1}\phi_{k_2}\cdots\phi_{k_N}\rangle = -|\phi_{k_2}\phi_{k_1}\cdots\phi_{k_N}\rangle$. The dimension of this space is
 $C_K^N = \frac{K!}{N!(K-N)!}$. Any wavefunction in \mathcal{H}_N can be expanded as

$$|\Psi\rangle = \sum_{k_1 < k_2 < \cdots < k_N} |\phi_{k_1}\phi_{k_2}\cdots\phi_{k_N}\rangle C_{k_1 k_2 \cdots k_N}. \quad (57)$$

Occupation number representation

Example for $(K, N) = (4, 2)$

$$\mathcal{H}_{N=2} = \text{span}\{|12\rangle, |13\rangle, |14\rangle, |23\rangle, |24\rangle, |34\rangle\}, \quad \dim(\mathcal{H}_{N=2}) = C_K^N = 6 \quad (58)$$

Another way to label these Slater determinants is by their occupation patterns of the four spin-orbitals, $|n_1 n_2 n_3 n_4\rangle$, called **occupation number vectors** (ONV)

$$|12\rangle \leftrightarrow |1100\rangle, \quad |13\rangle \leftrightarrow |1010\rangle, \quad |14\rangle \leftrightarrow |1001\rangle, \quad (59)$$

$$|23\rangle \leftrightarrow |0110\rangle, \quad |24\rangle \leftrightarrow |0101\rangle, \quad |34\rangle \leftrightarrow |0011\rangle. \quad (60)$$

Generalizing this idea, the expansion of $|\Psi\rangle$ can be written as

$$|\Psi\rangle = \sum_{\{n_k\}} |n_1 n_2 \cdots n_K\rangle C_{n_1 n_2 \cdots n_K}, \quad \sum_{k=1}^K n_k = N. \quad (61)$$

The language of second quantization

Second quantization originates from quantum field theory (QFT), in which the matter field $\psi(x)$ is also quantized,

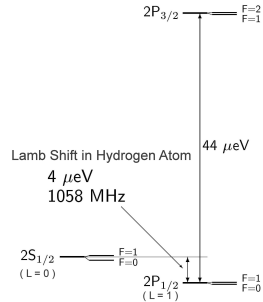
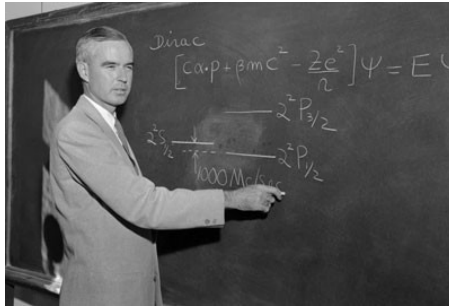
$$\hat{\psi}(x) = \sum_k \phi_k(x) a_k, \quad \hat{\psi}^\dagger(x) = \sum_k \phi_k(x) a_k^\dagger, \quad (62)$$

as well as the electromagnetic field.

Lamb shift: In 1947, Lamb and Retherford found from the hydrogen microwave spectrum that the $2S_{1/2}$ level is about 1 GHz (0.03 cm^{-1}) above the $2P_{1/2}$ level, which is inconsistent with the Dirac equation. It was the harbinger of modern quantum electrodynamics. This particular difference is a one-loop effect of quantum electrodynamics (QED), and can be interpreted as the influence of virtual photons that have been emitted and re-absorbed by the atom. For details, see <http://hyperphysics.phy-astr.gsu.edu/hbase/quantum/lamb.html#c3>
https://en.wikipedia.org/wiki/Lamb_shift
<https://ncatlab.org/nlab/show/Lamb+shift>

Lamb shift

Lamb shift (1947): $E_{2s_{1/2}} - E_{2p_{1/2}} = O(\alpha^3)$ [vs. relativistic effects $O(\alpha^{2n})$]



Another experiment (Kusch and Foley, 1947): anomalous magnetic moment of the electron $g_e > 2$. The current experimental value and uncertainty:

$a_e = \frac{g_e - 2}{2} = 0.001\,159\,652\,180\,73(28)$. Lamb and Foley were awarded Nobel Prize in Physics in 1955.

The language of second quantization

Second quantization turns out to be a very useful mathematical tool for the development of **quantum many-body theory** for condense matter physics and quantum chemistry.

The advantage of second quantization is that it provides a systematic way for the following tasks:

- ① Expressing wavefunction ansatz;
- ② Algebraic manipulation of equations;
- ③ Evaluation of matrix elements by Wick's theorem.

We will introduce the basics for (1) and (2), while (3) is more technical and will only be briefly mentioned.

Rules of second quantization

Without going into the details of justification, we just summarize the rules of second quantization necessary for our purpose.

Rule 1: Creation/Annihilation operators

$$a_k^\dagger |n_1 n_2 \cdots n_K\rangle = (1 - n_k)(-1)^{\sum_{i=1}^{k-1} n_i} |n_1 n_2 \cdots 1_k \cdots n_K\rangle, \quad (63)$$

$$a_k |n_1 n_2 \cdots n_K\rangle = n_k(-1)^{\sum_{i=1}^{k-1} n_i} |n_1 n_2 \cdots 0_k \cdots n_K\rangle. \quad (64)$$

Rule 2: Anti-commutation relations for Fermionic operators

$$\{a_i^\dagger, a_j^\dagger\} = \{a_i, a_j\} = 0, \quad \{a_i^\dagger, a_j\} = \delta_{ij}. \quad (65)$$

where the anticommutator is defined as $\{A, B\} = AB + BA$.

Remarks: These rules are consistent with the **antisymmetry principle**.

Rules of second quantization

Rule 3: Representation of operators

One-body operators $H_1 = \sum_{i=1}^N h(x_i)$:

$$\hat{H}_1 = \int dx \hat{\psi}^\dagger(x) h(x) \hat{\psi}(x) \quad (66)$$

or more explicitly,

$$\hat{H}_1 = \sum_{pq} \langle \phi_p | h | \phi_q \rangle a_p^\dagger a_q, \quad \langle \phi_p | h | \phi_q \rangle = \int dx \phi_p^\dagger(x) h(x) \phi_q(x) \quad (67)$$

Rules of second quantization

Rule 3: Representation of operators

Two-body operators $H_2 = \sum_{i < j}^N g(x_i, x_j)$ with $g(x_i, x_j) = g(x_j, x_i)$:

$$\hat{H}_2 = \frac{1}{2} \int dx_1 dx_2 \hat{\psi}^\dagger(x_1) \hat{\psi}^\dagger(x_2) g(x_1, x_2) \hat{\psi}(x_2) \hat{\psi}(x_1). \quad (68)$$

or more explicitly,

$$\hat{H}_2 = \frac{1}{2} \sum_{pqrs} \langle \phi_p \phi_q | g | \phi_r \phi_s \rangle a_p^\dagger a_q^\dagger a_s a_r, \quad (69)$$

$$\langle \phi_p \phi_q | g | \phi_r \phi_s \rangle = \int dx_1 dx_2 \phi_p^\dagger(x_1) \phi_q^\dagger(x_2) g(x_1, x_2) \phi_r(x_1) \phi_s(x_2). \quad (70)$$

The two-electron integral $\langle \phi_p \phi_q | g | \phi_r \phi_s \rangle$ is usually abbreviated as $\langle pq | rs \rangle$ (physicists' notation) or $[pr | qs]$ (chemists' notation).

Application 1: Action on states

(1) Consider the action of $a_3^\dagger a_1$ and $a_3^\dagger a_2$ on $|1100\rangle$:

$$a_3^\dagger a_1 |1100\rangle = a_3^\dagger (|0100\rangle) = -|0110\rangle, \quad (71)$$

$$a_3^\dagger a_2 |1100\rangle = a_3^\dagger (-|1000\rangle) = |1010\rangle. \quad (72)$$

The results are consistent with the replacement operators \hat{X}_{kl} acting on Slater determinants $\hat{X}_{31}|12\rangle = |32\rangle = -|23\rangle$ and $\hat{X}_{32}|12\rangle = |13\rangle$. Thus, we will just refer $a_k^\dagger a_l$ ($k \neq l$) as **replacement/excitation operator**.

(2) Consider the action of $a_3^\dagger a_3$ and $a_1^\dagger a_1$,

$$a_3^\dagger a_3 |1100\rangle = 0, \quad a_1^\dagger a_1 |1100\rangle = |1100\rangle. \quad (73)$$

Thus, $a_k^\dagger a_k = n_k$ will be referred as occupation number operator. In particular, $\hat{N} = \sum_{k=1}^K n_k$ is the **particle number operator**, viz., $\hat{N}|1100\rangle = 2|1100\rangle$.

Application 2: Evaluation of matrix elements

Slater-Condon rule for matrix elements over Slater determinants

$$\langle \Phi | \hat{H} | \Phi \rangle = \sum_i^N h_{ii} + \sum_{i < j}^N \langle ij || ij \rangle \quad (74)$$

$$\langle \Phi | \hat{H} | \Phi_i^a \rangle = h_{ia} + \sum_j^N \langle ij || aj \rangle \quad (75)$$

$$\langle \Phi | \hat{H} | \Phi_{ij}^{ab} \rangle = \langle ij || ab \rangle \quad (76)$$

$$\langle \Phi | \hat{H} | \Phi_{ijk}^{abc} \rangle = 0 \quad (77)$$

Here, we show that the second quantization can be used to derive these results in an algebraic way.

Application 2: Evaluation of matrix elements

For simplicity, we just examine the following cases for one-body operators:

Case (1): $\langle 12 | \hat{H}_1 | 12 \rangle = h_{11} + h_{22}$

$$\begin{aligned}\langle 12 | \hat{H}_1 | 12 \rangle &= \langle vac | a_2 a_1 \hat{H}_1 a_1^\dagger a_2^\dagger | vac \rangle = \sum_{pq} \langle vac | a_2 a_1 h_{pq} a_p^\dagger a_q a_1^\dagger a_2^\dagger | vac \rangle \\ &= \sum_{pq} \langle vac | a_2 a_1 h_{pq} a_p^\dagger (\{a_q, a_1^\dagger\} - a_1^\dagger a_q) a_2^\dagger | vac \rangle \\ &= \sum_p \langle vac | a_2 a_1 h_{p1} a_p^\dagger a_2^\dagger | vac \rangle - \sum_{pq} \langle vac | a_2 a_1 h_{pq} a_p^\dagger a_1^\dagger a_q a_2^\dagger | vac \rangle\end{aligned}$$

The first term gives h_{11} , while applying the anti-commutation relation again to the second term will leads to h_{22} .

$$\begin{aligned}\sum_{pq} \langle vac | a_2 a_1 h_{pq} a_p^\dagger a_1^\dagger \{a_q, a_2^\dagger\} | vac \rangle &= \sum_p \langle vac | a_2 a_1 h_{p2} a_p^\dagger a_1^\dagger | vac \rangle \\ &= \langle vac | a_2 h_{12} a_1^\dagger | vac \rangle - \sum_p \langle vac | a_2 h_{p2} a_p^\dagger a_1 a_1^\dagger | vac \rangle = -h_{22}\end{aligned}\tag{78}$$

Application 2: Evaluation of matrix elements

Case (2): $\langle 12|\hat{H}_1|13\rangle = h_{23}$

$$\begin{aligned}\langle 12|\hat{H}_1|13\rangle &= \langle vac|a_2a_1\hat{H}_1a_1^\dagger a_3^\dagger|vac\rangle \\ &= \sum_p \langle vac|a_2a_1h_{p1}a_p^\dagger a_3^\dagger|vac\rangle - \sum_{pq} \langle vac|a_2a_1h_{pq}a_p^\dagger a_1^\dagger a_q a_3^\dagger|vac\rangle.\end{aligned}$$

In this case, the first term vanishes, while the second becomes

$$\begin{aligned}\sum_{pq} \langle vac|a_2a_1h_{pq}a_p^\dagger a_1^\dagger a_q a_3^\dagger|vac\rangle &= \sum_{pq} \langle vac|a_2a_1h_{pq}a_p^\dagger a_1^\dagger \{a_q, a_3^\dagger\}|vac\rangle \\ &= \sum_p \langle vac|a_2a_1h_{p3}a_p^\dagger a_1^\dagger|vac\rangle = \sum_p (\langle vac|a_2h_{13}a_1^\dagger|vac\rangle - \langle vac|a_2h_{p3}a_p^\dagger a_1 a_1^\dagger|vac\rangle) \\ &= \sum_p -\langle vac|a_2h_{p3}a_p^\dagger a_1 a_1^\dagger|vac\rangle = \sum_p -\langle vac|a_2h_{p3}a_p^\dagger|vac\rangle = -h_{23}\end{aligned}\tag{79}$$

Application 2: Evaluation of matrix elements

Case (3): $\langle 12|\hat{H}_1|34\rangle = 0$

$$\begin{aligned}\langle 12|\hat{H}_1|34\rangle &= \langle vac|a_2a_1\hat{H}_1a_3^\dagger a_4^\dagger|vac\rangle \\ &= \sum_p \langle vac|a_2a_1h_{p3}a_p^\dagger a_4^\dagger|vac\rangle - \sum_{pq} \langle vac|a_2a_1h_{pq}a_p^\dagger a_3^\dagger a_q a_4^\dagger|vac\rangle.\end{aligned}$$

Thus, the matrix element is zero, since both terms vanish.

These are very simple illustration of the use of second quantization in evaluation of matrix elements. In general, we can derive the full set of Slater-Condon rules just using second quantization with anticommutation relation.

Application 2: Evaluation of matrix elements

However, it is also clear that using the anti-commutation relation becomes tedious as the number of operators increases. Luckily, there is a theorem named **Wick's theorem** for computing such matrix elements over physical vacuum (or more generally, Fermi vacuum after particle-hole transformation) systematically,

$$\langle vac | a_{p_1}^{(\dagger)} a_{p_2}^{(\dagger)} \cdots a_{p_k}^{(\dagger)} | vac \rangle \quad (80)$$

by finding all possible full pairwise contractions.

More powerfully, in combination with permutation symmetries of integrals, **diagrammatic** tools have been developed, which enables very simple identification of nonzero matrix elements. These topics, belonging to quantum many-body theory, will not be covered in this course, see Refs.²³

²[Ingvar Lindgren and John Morrison.](#) *Atomic many-body theory.* Springer-Verlag Berlin Heidelberg, 1982.

³[Isaiah Shavitt and Rodney J Bartlett.](#) *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory.* Cambridge university press, 2009.

Application 3: Spin operators

In many-electron systems, the total spin angular momentum operator is $\hat{S}_\alpha = \sum_{i=1}^N \hat{s}_\alpha(i)$ ($\alpha = x, y, z$). The spin square operator is

$$\hat{S}^2 = \hat{S}_x^2 + \hat{S}_y^2 + \hat{S}_z^2 = \hat{S}_+ \hat{S}_- - \hat{S}_z + \hat{S}_z^2 = \hat{S}_- \hat{S}_+ + \hat{S}_z + \hat{S}_z^2 \quad (81)$$

In second quantization, the spin angular momentum operators are⁴

$$\hat{S}_+ = \sum_p a_{p\alpha}^\dagger a_{p\beta} \quad (82)$$

$$\hat{S}_- = \sum_p a_{p\beta}^\dagger a_{p\alpha} = \hat{S}_+^\dagger \quad (83)$$

$$\hat{S}_z = \frac{1}{2} \sum_p (a_{p\alpha}^\dagger a_{p\alpha} - a_{p\beta}^\dagger a_{p\beta}) = \frac{1}{2} (\hat{N}_\alpha - \hat{N}_\beta) \quad (84)$$

For nonrelativistic (more precisely, spin-free) \hat{H} , $(\hat{H}, \hat{N}, \hat{S}^2, \hat{S}_z)$ commute with each other, eigenfunctions of \hat{H} can be labelled by quantum numbers (N, S, M_S) .

⁴Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.

Application 3: Spin operators

We can use the spin operators in second quantized form to check whether a given state is a spin eigenfunction.

Example 1: $|\Phi\rangle = |\bar{i}\bar{i}\rangle$ (singlet)

$$\begin{aligned}\hat{S}_z |\Phi\rangle &= \frac{1}{2}(a_{i\alpha}^\dagger a_{i\alpha} - a_{i\beta}^\dagger a_{i\beta}) |\bar{i}\bar{i}\rangle = \frac{1}{2}(|\bar{i}\bar{i}\rangle - |\bar{i}\bar{i}\rangle) = 0 \\ \hat{S}_+ \hat{S}_- |\Phi\rangle &= (a_{i\alpha}^\dagger a_{i\beta})(a_{i\beta}^\dagger a_{i\alpha}) |\bar{i}\bar{i}\rangle = a_{i\alpha}^\dagger a_{i\beta} a_{i\beta}^\dagger a_{i\alpha} |\bar{i}\bar{i}\rangle = 0 \\ \hat{S}^2 |\Phi\rangle &= (\hat{S}_+ \hat{S}_- - \hat{S}_z + \hat{S}_z^2) |\bar{i}\bar{i}\rangle = 0\end{aligned}\tag{85}$$

Example 2: $|\Phi\rangle = |ij\rangle$ (triplet, $M_S = 1$)

$$\begin{aligned}\hat{S}_z |\Phi\rangle &= \sum_k \frac{1}{2}(a_{k\alpha}^\dagger a_{k\alpha} - a_{k\beta}^\dagger a_{k\beta}) |ij\rangle = \frac{1}{2}(|ij\rangle + |ij\rangle) = |ij\rangle \\ \hat{S}_- \hat{S}_+ |\Phi\rangle &= \hat{S}_- \hat{S}_+ |ij\rangle = 0 \\ \hat{S}^2 |\Phi\rangle &= (\hat{S}_- \hat{S}_+ + \hat{S}_z + \hat{S}_z^2) |ij\rangle = (0 + |ij\rangle + |ij\rangle) = 2 |ij\rangle\end{aligned}\tag{86}$$

Application 3: Spin operators

Now let us introduce spin tensor operators:⁵

Single excitation operators

$$\hat{S}_{pq}(0,0) = \frac{1}{\sqrt{2}}(a_{p\alpha}^\dagger a_{q\alpha} + a_{p\beta}^\dagger a_{q\beta}) \quad (87)$$

$$\hat{T}_{pq}(1,1) = -a_{p\alpha}^\dagger a_{q\beta} \quad (88)$$

$$\hat{T}_{pq}(1,0) = \frac{1}{\sqrt{2}}(a_{p\alpha}^\dagger a_{q\alpha} - a_{p\beta}^\dagger a_{q\beta}) \quad (89)$$

$$\hat{T}_{pq}(1,-1) = a_{p\beta}^\dagger a_{q\alpha} \quad (90)$$

These operators generate singlet and triplet excited states acting on a close-shell reference. In particular, $|\Phi_\pm\rangle = \frac{1}{\sqrt{2}}(a_{a\alpha}^\dagger a_{i\alpha} \pm a_{a\beta}^\dagger a_{i\beta})|i\bar{i}\rangle = \frac{1}{\sqrt{2}}(|a\bar{i}\rangle \pm |i\bar{a}\rangle)$ is singlet (triplet) for $+$ ($-$).

⁵Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.

Application 3: Spin operators

For simplicity, we examine the expectation value

$$\begin{aligned}\langle \Phi_{\pm} | \hat{S}^2 | \Phi_{\pm} \rangle &= \langle \Phi_{\pm} | (\hat{S}_- \hat{S}_+ + \hat{S}_z + \hat{S}_z^2) | \Phi_{\pm} \rangle \\ &= \langle \Phi_{\pm} | \hat{S}_- \hat{S}_+ | \Phi_{\pm} \rangle = \langle \hat{S}_+ \Phi_{\pm} | \hat{S}_+ \Phi_{\pm} \rangle\end{aligned}\quad (91)$$

The state $|\hat{S}_+ \Phi_{\pm}\rangle$ is simply

$$|\hat{S}_+ \Phi_{\pm}\rangle = \hat{S}_+ \frac{1}{\sqrt{2}}(|a\bar{i}\rangle \pm |i\bar{a}\rangle) = \frac{1}{\sqrt{2}}(|ai\rangle \pm |ia\rangle) = \frac{1}{\sqrt{2}}(|ai\rangle \mp |ai\rangle) \quad (92)$$

Therefore, $\langle \Phi_{\pm} | \hat{S}^2 | \Phi_{\pm} \rangle = \langle \hat{S}_+ \Phi_{\pm} | \hat{S}_+ \Phi_{\pm} \rangle$ is 0 (singlet) or 2 (triplet).

Singlet and triplet states with two open-shell electrons

$$|\Phi_{\pm}\rangle = \frac{1}{\sqrt{2}}(a_{a\alpha}^{\dagger} a_{i\alpha} \pm a_{a\beta}^{\dagger} a_{i\beta}) |i\bar{i}\rangle = \frac{1}{\sqrt{2}}(|a\bar{i}\rangle \pm |i\bar{a}\rangle) \quad (93)$$

Reference I

- [1] Ernest R Davidson. “The iterative calculation of a few of the lowest eigenvalues and corresponding eigenvectors of large real-symmetric matrices”. In: *Journal of Computational Physics* 17 (1975), pp. 87–94.
- [2] Paul Adrien Maurice Dirac. “Quantum mechanics of many-electron systems”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733.
- [3] Trygve Helgaker, Poul Jorgensen, and Jeppe Olsen. *Molecular electronic-structure theory*. John Wiley & Sons, 2014.
- [4] Ingvar Lindgren and John Morrison. *Atomic many-body theory*. Springer-Verlag Berlin Heidelberg, 1982.
- [5] Isaiah Shavitt and Rodney J Bartlett. *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*. Cambridge university press, 2009.