NTNU
Norwegian University of
Science and Technology

Symmetric Coupled Cluster Theory

Eirik Fadum Kjønstad

ABSTRACT. We introduce a novel coupled cluster formulation for the ground state, symmetric coupled cluster (SCC) theory, and an extension to excited states, molecular properties, and transition elements called symmetric equation of motion (SEOM) theory. The ground state cluster amplitudes and the expansion coefficients of the excited states are determined by stationarity of a Hermitian approximation of the Hamiltonian expectation value.

We present an implementation of the ground state energy at the singles and doubles (SCCSD) truncation of the cluster operator. A series of calculations on molecular systems comprised of fewer than 50 electrons demonstrate that the SCCSD and CCSD energies are of similar accuracy. The method does not scale correctly with system size, and the errors introduced as a consequence limit its applicability to small systems.

SAMMENDRAG. Vi introduserer en modell for grunntilstanden, symmetrisk coupled-cluster-teori (SCC), samt en utvidelse til eksiterte tilstander, molekylegenskaper og overgangselementer kalt symmetrisk equation-of-motion-teori (SEOM). Grunntilstandens amplityder og de eksiterte tilstanders ekspansjonskoeffisienter bestemmes ved stasjonæritet av en Hermitesk tilnærming av Hamilton-operatorens forventningsverdi.

En implementering av grunntilstandsenergien for en singles- og doubles-trunkering av cluster-operatoren (SCCSD) presenteres. Beregninger på en rekke molekylære systemer bestående av færre enn 50 elektroner viser at SCCSD- og CCSD-energier har lignende nøyaktighet. Metodens feilaktige skalering med systemets størrelse medfører at den bare kan benyttes på små systemer.

Preface

The work presented was conducted at the Norwegian University of Science and Technology for the Department of Chemistry and concludes the five-year master's degree programme "Chemical Engineering and Biotechnology" with specialization in applied theoretical chemistry. The master's thesis was supervised by Prof. Henrik Koch and written during the spring of 2016.

Parts of the research was done during my visit to the group of Prof. Todd J. Martínez at Stanford University Jan–Feb 2016. The work on both theory and implementation was carried out in collaboration with Prof. Koch.

Contents

List of figures

List of tables

Chapter 1

Introduction

Today, coupled cluster theory is the most accurate practical model for the description of instantaneous electron-electron interactions not accounted for in a mean-field approximation.[1] Its modern variant, with the exponential *ansatz* for wave function and subsequent projection of Schrödinger's equation, was first introduced in the late 1950s by Coester and Kümmel[2,3] and soon after brought into quantum chemistry by Čížek and Paldus.[4,5] Coupled cluster theory has been in wide use since the 1980s, when the first successful implementations were reported at the singles and doubles (CCSD)[6] as well as triples (CCSDT)[7] levels of theory. A hallmark feature of the coupled cluster model is its correct scaling with the system size, a property some authors have emphasized to explain its superior performance[8] to configuration interaction[9] theory. Although it is to date preferable to treat large molecular systems with the cheaper and less accurate second-order Møller-Plesset[10] (MP2) and density functional theories,[11] coupled cluster theory is increasingly applied on bigger systems, with current extensive use for molecules comprised of up to about 15 first- and second-row atoms.[12]

The description of time-dependent phenomena in coupled cluster theory was introduced by Monkhorst[13] and Dalgaard and Monkhorst,[14] who made use of response theory,[15,16] *i.e.* perturbation theory,[17] to time-evolve suitably defined expectation values. After the introduction of the correct expectation value,[18] others later derived the general linear and quadratic[19] and cubic[20] coupled cluster response functions. These funtions encode the expansion of expectation values in orders of the perturbation—the nth order molecular properties—and the excitation energies and transition matrix elements may be derived from their poles and residues. Response theory has been utilized in a range of quantum chemical models, see the recent and comprehensive review by Helgaker *et al.*[21] and references therein.

Bartlett and coworkers have developed an alternative approach, equation of motion coupled cluster theory.[22] In this approach the reference system's excited states are constructed, and properties and transition moments are derived by substitution in the spectral representation of the linear response function. As a time-independent formalism, it is difficult to obtain higher-order molecular properties, see Rozynczko and Bartlett[23,24] and correspondence.[25,26] A notable feature of the theory is the explicit construction of the states, allowing extensions to describe ionization[27] and electron attachment[28]

processes. The expansion of the states, which bears resemblance to configuration interaction theory, moreover endows it with the flexibility to describe multireference character in the excited states.[29]

A major drawback of ground state coupled cluster theory is its inability to properly describe chemical systems whose true wavefunction contains several leading determinants. The exponential state has a large number of small contributions along determinants other than the reference, thereby accounting for instantaneuous interactions or dynamical correlation, but most often the determinant of highest weight is the Hartree-Fock[30] determinant.[31] The theory's accuracy deteriorates as a consequence whenever the reference provides a poor zeroth order description of the wavefunction.[32] This multireference regime is ubiquitous in chemistry. It is frequently encountered in bond breaking/formation processes, open-shell and excited electronic states, transition-metal complexes, and on occasion at the equilibrium geometry of the ground state.[31] The formulation of multireference coupled cluster theories has turned out to be very challenging, with no methods in wide use at the time of writing. The reader is referred to the recent review by Lyakh *et al.*[31] More practical at present is the use of multiconfigurational self-consistent field theory, which may provide an adequate reference upon which multireference configuration interaction models can lead to quantitative results.[33]

Electronic degeneracies, and in particular conical intersections,[34] are now widely recognized as a common feature in the photochemistry of organic molecules.[35–37] In terms of their geometry, such intersections assume the shape of two facing cones[38] and serve as funnels for nuclear motion by providing radiation-less paths between potential energy surfaces.[39] The accurate description of conical intersections has in recent years become a priority for quantum chemical methods, with ongoing efforts to assess the performance of established theories.[40–44] In the vicinity of electronic degeneracies, a first-principles approach to dynamics becomes vital, as quantum chemistry provides the sole means to estimate the non-adiabatic coupling matrix elements that influences the nuclear dynamics.[36] Notable in this regard is the *ab initio* multiple spawning theory of Martínez and coworkers.[45,46] Recent developments in ultra-fast pump-probe experimental techniques have made such models amenable to complementary experimental investigation.[47]

Some time ago, Hättig questioned whether coupled cluster theories are able to describe a conical intersection.[48] An analysis of non-Hermitian eigenvalue equations suggested that the obtained dimension of the intersection is wrong, and that complex excitation energies could appear in its vicinity. The occurence of complex energies was later observed at an intersection in formaldehyde.[49] As a consequence, several authors[40,48,50] have advocated the Hermitian, size-intensive[51] and comparably accurate[52] algebraic diagrammatic construction (ADC) method of Schirmer.[53] The hierarchy of ADC(n) models are derived from a perturbation expansion of the linear response

function. As such, they are only applicable when MPn provides an adequate zeroth order description of the wavefunction,[50] making them less robust than coupled cluster theories, which offer a superior treatment of dynamic correlation[1] and approximate orbital invariance.[54] However, as pointed out by Hättig,[48] the ADC(2) method is closely related to the perturbative CC2[55] coupled cluster model.

Coupled cluster theory's non-Hermiticity, in its traditional formulation, is well known to stem from the projection principle.[22,49,56,57] This loss of symmetry is responsible for the non-variationality of the energy,[58] for instance causing a catastrophic collapse in N_2.[59] It breaks certain symmetries of the response functions, leading to inconsistencies in molecular properties and endowing origin dependence to quantities independent of the origin.[57] It is moreover responsible for the non-orthogonality of equation of motion's excited states, necessitating the construction of a set of 'dual states' to evaluate transition moments.[22] A dual state is also introduced to evaluate ground state energy gradients and response functions, as it is necessary for a generalized Hellmann-Feynman[60] theorem to be valid.[19]

A common feature in attempted symmetric formulations of coupled cluster theory is the reinstatement of variationality by eliminating projection.[28] This approach unfortunately leads to a lack of finiteness: the energy and the equations that determine the wavefunction become infinite series.[61] In addition to this elimination, Taube and Bartlett[62] have commented on the advantages of a unitary exponential *ansatz* (which dates back to Yaris[63]), noting in particular its ability to unify the response and equation of motion formalisms. Due to their non-finiteness, these approaches must employ truncation schemes, introducing both assumptions on the convergence of the series (*e.g.* that the cluster operator is small) and sacrificing their variational nature. The central trade-off encountered in non-projective theories was effectively summed up twenty years ago by Szalay, Nooijen, and Bartlett, when they concluded that "attempts to introduce *finite* and *symmetric* [formulations] also associated with the satisfaction of the [generalized Hellmann-Feynman] theorem have not been successful [. . .] we are rather pessimistic whether such a formulation can exist."[56]

We present in this thesis a singles and doubles implementation of a novel coupled cluster formulation, symmetric coupled cluster (SCC) theory, which is non-variational, symmetric, finite, and consistent with a generalized Hellmann-Feynman theorem. It is best understood as an attempt to introduce the projection principle in a symmetric fashion. Owing to the theory's projective flavor it bears close resemblance to traditional coupled cluster theory, providing the same solutions under certain circumstances. Its major caveat is its improper scaling with system size, limiting its applicability to small molecular systems.

A response theoretical approach to time-dependent phenomena is prefer-

able in our view, as it defines properties to all orders and its response functions and transition elements scale correctly with size.[64,65] In this preliminary study, due to its ease of derivation, we present the symmetric equation of motion (SEOM) extension of our ground state model. The eigenvalue problem for the excitation energies as well as the expressions for the first-order molecular properties and transition moments are derived.

As theoretical chemists we attempt to formulate the perfect theory: accurate, extensive, symmetric, finite, variational, and potentially exact. In reality, the fulfillment of one desirable theoretical property is often seen to go hand-in-hand with the loss of another. The theory may either be finite or Hermitian; symmetric or extensive; gauge invariant or exact in the limit. Valuable lessons may often be drawn from such encounters, even if the models never reach maturity and widespread use.

Chapter 2

Selected topics in quantum chemistry

2.1. FOUNDATIONS

The behavior of a physical system is, by the tenets of quantum theory, described in terms of its quantum state Ψ, an object that evolves in time according to the celebrated Schrödinger equation[66]

$$i\,\frac{\partial \Psi}{\partial t} = H\,\Psi. \tag{2.1}$$

At each instant in time, the state Ψ inhabits the Hilbert space $\mathscr{H}$ of square integrable functions.[67] Systems that do not depend explicitly on time are described by the quantum states $\Psi_n(t) = \psi_n\,e^{-iE_n t}$, where[68]

$$H\,\psi_n = E_n\,\psi_n. \tag{2.2}$$

This is the time-independent Schrödinger equation, solved to determine the wavefunctions ψ_n and energies E_n of the physical system governed by the Hamiltonian operator H.

Unless otherwise stated, we will throughout the text consider H to be the clamped-nuclei or molecular electronic Hamiltonian, describing a chemical system with the nuclei fixed in a particular molecular geometry. It can be written, in atomic units, as[69]

$$H = -\frac{1}{2}\sum_i \boldsymbol{\nabla}_i^2 - \sum_i\sum_I \frac{Z_I}{|\mathbf{r}_i - \mathbf{R}_I|} + \sum_i\sum_{j>i} \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|}, \tag{2.3}$$

where $\mathbf{r}_i$ and $\mathbf{R}_I$ refers to the position of electron i and nucleus I, respectively, and Z_I is the charge of the Ith nucleus.

2.1.1. The Fock space approximation

If the electrons in a molecule move independently of one another, the wavefunction becomes a product of one-particle functions. Anti-symmetrizing this wavefunction so as to be in agreement with the Pauli principle[69] results in the *Slater determinant* (hereafter simply *determinant*)

$$\psi(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \ldots & \varphi_1(\mathbf{r}_N) \\ \varphi_2(\mathbf{r}_1) & \varphi_2(\mathbf{r}_2) & \ldots & \varphi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_N(\mathbf{r}_1) & \varphi_N(\mathbf{r}_2) & \ldots & \varphi_N(\mathbf{r}_N) \end{vmatrix}, \tag{2.4}$$

where the φ_i are termed *spin orbitals* and N is the number of electrons in the system. For brevity, we have suppressed the spin coordinate s_i of the ith electron, *i.e.*, $\mathbf{r}_i \equiv \mathbf{r}_i, s_i$, where $\mathbf{r}_i$ refers to its spatial position. In non-relativistic theories, the one-electron spin orbitals φ_i are commonly written $\varphi_i(\mathbf{r}_i, s_i) = \vartheta_i(\mathbf{r}_i)\,\sigma_i(s_i)$, where ϑ_i is called a *molecular orbital* and σ_i is an eigenfunction of the spin operator.[1]

Approximate wavefunctions are often constructed as linear combinations of determinants. For a molecular system of N electrons, we define the (N-electron) *Fock space* $\mathscr{F}$ as the space spanned by all N-particle determinants that may be formed from an assumed finite and given orthonormal spin orbital basis $\{\varphi_i\}_{i=1}^M$. Denoting the basis of determinants as $\mathcal{M} = \{\Phi_\mu\}_\mu$, the Schrödinger equation may be solved within the subspace $\mathscr{F} \subset \mathscr{H}$,

$$\mathbf{H}\,\mathbf{c}_n = E_n\,\mathbf{c}_n, \tag{2.5}$$

where $H_{\mu\nu} = \langle \Phi_\mu, H\,\Phi_\nu \rangle$ and $\mathbf{c}_n = (c_{nm})_m$ contains the expansion coefficients of the nth eigenstate $\psi_n = \sum_m c_{nm}\,\Phi_m$. Henceforth we shall by H mean the operator on $\mathscr{F}$ associated with the matrix $\mathbf{H}$. This is what we might refer to as the *Fock space approximation*, appropriate only if the chosen spin orbital basis $\{\varphi_i\}_i$ offers a sufficiently accurate description of the chemical system under study. In the remainder of the text we will for the most part adopt Dirac's bracket notation,[70] writing ψ as $|\psi\rangle$ and inner products as $\langle \varphi, \psi \rangle = \langle \varphi \,|\, \psi \rangle$ and $\langle \varphi, A\,\psi \rangle = \langle \varphi \,|\, A \,|\, \psi \rangle$.

2.1.2. The variational principle

The variational principle directs us to select the wavefunction $|\psi\rangle$ which minimizes the expectation value $\langle H \rangle$ of H. Its rationale presents itself from the useful variational theorem[67]

$$\inf \sigma(H) \leq \frac{\langle \psi \,|\, H \,|\, \psi \rangle}{\langle \psi \,|\, \psi \rangle} \leq \sup \sigma(H), \tag{2.6}$$

where $\sigma(H) = \{E_n\}_n$ is called the *spectrum*[71] of H. As $\langle H \rangle$ always exceeds the ground state energy $E_0 = \inf \sigma(H)$, it is clearly reasonable to select the wavefunction that minimizes $\langle H \rangle$. The variational principle represents a versatile tool to identify accurate approximations of the ground state.

A particularly important case is the minimization of $\langle H \rangle$ with respect to the spin orbitals $\{\varphi_i\}_i$ of a single determinant. The obtained solution is the well known Hartree-Fock state $|\mathrm{R}\rangle$.[30] We assume throughout the text that the set $\{\varphi_i\}_{i=1}^M$ is ordered such that $|\mathrm{R}\rangle$ is formed from the N first orbitals:

$$|\mathrm{R}\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \varphi_1(\mathbf{r}_1) & \varphi_1(\mathbf{r}_2) & \ldots & \varphi_1(\mathbf{r}_N) \\ \varphi_2(\mathbf{r}_1) & \varphi_2(\mathbf{r}_2) & \ldots & \varphi_2(\mathbf{r}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_N(\mathbf{r}_1) & \varphi_N(\mathbf{r}_2) & \ldots & \varphi_N(\mathbf{r}_N) \end{vmatrix}. \tag{2.7}$$

The orbitals φ_i with $i = 1, 2, \ldots, N$ are called *occupied* whereas the φ_a with $a = N+1, \ldots, M$ are referred to as *virtual*. We reserve the letters $i, j, k, l, \ldots$ for occupied orbitals and $a, b, c, d, \ldots$ for virtual orbitals.

The reader is assumed to be familiar with the second quantization representation of quantum mechanics, of which many good treatments can be found elsewhere.[15,16,21,72] In this formalism, particles are placed in orbitals φ_i by creation operators $a_i^\dagger$ and removed from them by annihilation operators a_i. These operators satisfy the anti-commutator relations

$$[a_p^\dagger, a_q^\dagger]_+ = 0, \tag{2.8a}$$

$$[a_p, a_q]_+ = 0, \tag{2.8b}$$

$$[a_p^\dagger, a_q]_+ = \delta_{pq}. \tag{2.8c}$$

The reference may be written $|\mathrm{R}\rangle = (\prod_{i=1}^N a_i^\dagger) |\mathrm{vac}\rangle$, where the *vacuum state* $|\mathrm{vac}\rangle$ contains no electrons. One- and two-electron operators are written

$$A = \sum_{pq} A_{pq}\, a_p^\dagger a_q, \quad B = \frac{1}{2} \sum_{pqrs} B_{pqrs}\, a_p^\dagger a_r^\dagger a_s a_q, \tag{2.9}$$

respectively, where $A_{pq} = \langle \varphi_p, A\, \varphi_q \rangle$ and $B_{pqrs} = \langle \varphi_p\, \varphi_r, B\, \varphi_s\, \varphi_q \rangle$. We refer the reader to the excellent monograph by Helgaker *et al.*[1] for a self-contained treatment of the use of second quantization in chemistry.

It is useful to define a set of *excitation operators* τ_μ that exchange occupied and virtual orbitals, $\tau_0 = \mathbb{I}$, $\tau_i^a = a_a^\dagger a_i$, $\tau_{ij}^{ab} = \tau_j^b \tau_i^a$, and so on. The determinants in $\mathscr{F}$ are conveniently constructed by letting τ_μ act on the reference $|\mathrm{R}\rangle$:

$$\left| {a \atop i} \right\rangle = \tau_i^a |\mathrm{R}\rangle, \quad \left| {ab \atop ij} \right\rangle = \tau_{ij}^{ab} |\mathrm{R}\rangle, \quad \ldots \tag{2.10}$$

When the particular excitation operator τ_μ is unspecified, we denote the determinant as $|\mu\rangle$. These operators allow us to define a hierarchy of subspaces of the Fock space $\mathscr{F}$:

$$\mathscr{F}_0 = \mathrm{span}\, \mathcal{M}_0 = \mathrm{span}\, \{ |\mathrm{R}\rangle \}$$

$$\mathscr{F}_1 = \mathrm{span}\, \mathcal{M}_1 = \mathrm{span}\, \left\{ |\mathrm{R}\rangle, \left| {a \atop i} \right\rangle \right\}_{ai}$$

$$\mathscr{F}_2 = \mathrm{span}\, \mathcal{M}_2 = \mathrm{span}\, \left\{ |\mathrm{R}\rangle, \left| {a \atop i} \right\rangle, \left| {ab \atop ij} \right\rangle \right\}_{aibj} \tag{2.11}$$

$$\vdots$$

It may be noted that $\mathscr{F}_0 \subset \mathscr{F}_1 \subset \ldots \subset \mathscr{F}_N$, and that $\mathscr{F}_N = \mathscr{F}$, since no more than N electrons can be excited from the reference $|\mathrm{R}\rangle$.

Any wavefunction $|\psi\rangle \in \mathscr{F}_n$ may be written $|\psi\rangle = \sum_{\mu \geq 0} c_\mu |\mu\rangle$, where it is understood that the index μ is restricted to determinants in this space. Following the variational principle, we may now minimize $\langle H \rangle$ with respect to c_μ by locating its stationary points:

$$\frac{\partial}{\partial c_\mu} \frac{\langle \psi | H | \psi \rangle}{\langle \psi | \psi \rangle} = 0 \iff \mathbf{H}\,\mathbf{c} = E\,\mathbf{c}. \tag{2.12}$$

The matrix $\mathbf{H}$ is the representation of H in the basis $\mathcal{M}_n$ of $\mathscr{F}_n$. This matrix equation defines the nth order configuration interaction (CI) model.[9] In a full expansion ($n = N$) Equation (2.12) is equivalent to Equation (2.5), and the Fock space $\mathscr{F}$ is therefore sometimes called the full configuration interaction space.

The variational theorem tells us that the lowest eigenvalue of $\mathbf{H}$ is an upper bound to the exact ground state energy. However, an even stronger result exists, the Hylleraas-Undheim theorem.[73] It states that as the basis used in the matrix representation $\mathbf{H}$ is enlarged, the kth eigenvalue of $\mathbf{H}$ converges from above to the kth eigenvalue of H. In CI theory, $\mathcal{M}_0 \subset \mathcal{M}_1 \subset \ldots \subset \mathcal{M}_N = \mathcal{M}$, and the kth eigenvalue of $\mathbf{H}$ must be an upper bound to the true kth energy of the Hamiltonian H.

2.1.3. The Hellmann-Feynman theorem

Chemical systems are often best described in terms of a solvable time-independent reference system H_0 under the influence of a perturbation V. Introducing a strength parameter λ, this situation is encoded in the Hamiltonian $H = H_0 + \lambda V$. The theorem due to Hellmann and Feynman states that the first order change in $\langle H \rangle$ is equal to the expectation value of the perturbation V.[1] In particular, assuming that $|\psi\rangle$ is an eigenstate of H_0 prior to switching on the perturbation, the theorem reads

$$\frac{\mathrm{d}\langle H \rangle}{\mathrm{d}\lambda}\bigg|_{\lambda=0} = \langle \psi | V | \psi \rangle. \tag{2.13}$$

Apart from being a relation satisfied by exact states, its validity in approximate theories allows the expectation values of arbitrary operators $\mathcal{O}$ to be written in terms of energy derivatives:

$$\langle \mathcal{O} \rangle = \frac{\mathrm{d}\langle H \rangle}{\mathrm{d}\lambda}\bigg|_{\lambda=0}, \quad H = H_0 + \lambda \mathcal{O}. \tag{2.14}$$

2.1.4. Response theory for exact states

For a system perturbed by V, the powerful framework of perturbation theory[68] exploits the knowledge of the reference system H_0 to identify the exact solution in orders of the perturbation. In a nutshell, response theory

estimates the change in $\langle A \rangle$ due to V by perturbation theoretical considerations.[15] Perturbation theory is elegantly formulated in terms of *propagators*, unitary operators that propagate states from one point in time to another:

$$|\Psi(t)\rangle = U(t, t_0) |\Psi(t_0)\rangle. \tag{2.15}$$

A Schrödinger equation is readily formulated for the propagator. It is useful to introduce Dirac's interaction picture $|\Psi_I(t)\rangle = U_0(t, t_0)^\dagger |\Psi(t)\rangle$ in this endeavor. The associated propagator U_I may be shown to satisfy[74]

$$U_I(t, t_0) = \mathbb{I} - i \int_{t_0}^t V_I(t') \, U_I(t', t_0) \, \mathrm{d}t', \tag{2.16}$$

where $V_I(t) = U_0(t, t_0)^\dagger \, V(t) \, U_0(t, t_0)$. Perturbation theory is the iterative solution of this operator equation by repeated substitution:

$$U_I(t, t_0) = \mathbb{I} - i \int_{t_0}^t V_I(t') \, \mathrm{d}t' + i^2 \int_{t_0}^t \int_{t_0}^{t'} V_I(t') \, V_I(t'') \, \mathrm{d}t' \, \mathrm{d}t'' + \dots . \tag{2.17}$$

For a system in the state $|\Psi_0\rangle$ at t_0, perturbation theory predicts the system's time evolution by a truncation of the infinite series

$$|\Psi(t)\rangle = U_0(t, t_0)^\dagger \left(\mathbb{I} - i \int_{t_0}^t V_I(t') \, \mathrm{d}t' + \dots \right) |\Psi_0\rangle. \tag{2.18}$$

In response theory, it is not *per se* the time evolution of the states that is of interest. Rather, the evolution of the expectation values serve to define the central objects, the *response functions* $\langle\langle A; V^{\omega_1}, \dots, V^{\omega_n} \rangle\rangle_{\omega_1, \dots, \omega_n}$:

$$\langle A \rangle = \langle \Psi_0 | A | \Psi_0 \rangle + \int \mathrm{d}\omega \, e^{i\omega t} \langle\langle A; V^\omega \rangle\rangle_\omega$$
$$\iint \mathrm{d}\omega_1 \, \mathrm{d}\omega_2 \, e^{i(\omega_1 + \omega_2)t} \langle\langle A; V^{\omega_1}, V^{\omega_2} \rangle\rangle_{\omega_1, \omega_2} + \dots \tag{2.19}$$

They are derived by substituting Equation (2.18) in the definition of the expectation value, collecting the terms of equal order of V. The frequencies $\omega_1, \dots, \omega_n$ are introduced by the Fourier transform[75] $V = \int \mathrm{d}\omega \, e^{-i\omega t} V^\omega$. By resolving the identity as $\mathbb{I} = \sum_n |\psi_n\rangle\langle\psi_n|$ it is not difficult to derive

$$\langle\langle A; V^\omega \rangle\rangle_\omega = \sum_n \frac{\langle \psi_0 | A | \psi_n \rangle \langle \psi_n | V^\omega | \psi_0 \rangle}{\omega - \omega_n} -$$
$$\sum_n \frac{\langle \psi_0 | V^\omega | \psi_n \rangle \langle \psi_n | A | \psi_0 \rangle}{\omega + \omega_n}, \tag{2.20}$$

the spectral representation of the linear response function.[19] Its poles are located at the reference system's excitation energies ω_n and transition matrix elements $\Gamma_{n \to m}^{\mathcal{O}} = \langle \psi_m | \mathcal{O} | \psi_n \rangle$ are obtained from its residues:

$$\lim_{\omega \to \omega_k} (\omega - \omega_k) \langle\langle A; V^\omega \rangle\rangle_\omega = \Gamma_{0 \to k}^A \, \Gamma_{k \to 0}^{V^{\omega_k}}, \tag{2.21}$$

$$\lim_{\omega \to -\omega_k} (\omega + \omega_k) \langle\langle A; V^\omega \rangle\rangle_\omega = -\Gamma_{0 \to k}^{V^{-\omega_k}} \, \Gamma_{k \to 0}^A. \tag{2.22}$$

2.2. Coupled cluster theory

2.2.1. The *ansatz* and the amplitude equations

A detailed inspection of the perturbation series for the wavefunction $|\psi\rangle$ reveals that it may be written in an exponential form.[2,3] The coupled cluster *ansatz*[12] consequently posits that $|\psi\rangle$ and $|R\rangle$ are related as

$$|\psi\rangle = e^T |R\rangle, \tag{2.23}$$

where the *cluster operator* T is a sum of excitation operators τ_μ weighted by a set of *cluster amplitudes* t_μ. We write

$$T = \sum_\mu t_\mu \, \tau_\mu = \sum_{ai} t_i^a \, a_a^\dagger a_i + \sum_{ai \geq bj} t_{ij}^{ab} \, a_a^\dagger a_b^\dagger a_i a_j + \ldots \tag{2.24}$$

The amplitudes t_μ are determined by projection of the Schrödinger equation $H|\psi\rangle = E|\psi\rangle$ onto a basis of the subspace $\mathscr{F}_n \subset \mathscr{F}$. We will refer to this prescription to determine $|\psi\rangle$ as 'the projection principle' to distance it from the variational principle. By selecting the particular basis $\{\langle R|, \langle \mu| \, e^{-T}\}_\mu$ of $\mathscr{F}_n$ we obtain the *amplitude equations*:

$$E = \langle R| \, e^{-T} H \, e^T \, |R\rangle, \tag{2.25}$$

$$0 = \langle \mu| \, e^{-T} H \, e^T \, |R\rangle. \tag{2.26}$$

If the projection space contains up to n-fold excitations (*i.e.* it is $\mathscr{F}_n$) the cluster operator is similarly truncated as $T = T_1 + T_2 + \ldots + T_n$, where T_i includes only i-fold excitation operators. The order of truncation n defines the hierarchy of coupled cluster models: singles (CCS), singles and doubles[6] (CCSD), singles, doubles, and triples[7] (CCSDT), and so on. For practical purposes, the approximate inclusion of higher order excitations offer significant improvements, as is done in the CC2[55] and CC3[76] models (see also Koch *et al.*[77]) as well as the 'gold standard' CCSD(T) approximation.[78]

Consider a composite system consisting of two non-interacting systems A and B. In mathematical terms, this allows us to write the Hamiltonian as $H = H_A \otimes \mathbb{I} + \mathbb{I} \otimes H_B$, where H_A pertains to interactions within subsystem A and H_B within B. It is now straight-forward to show that $|\psi\rangle = |\psi_A\rangle \otimes |\psi_B\rangle$ is a solution of the Schrödinger equation with energy equal to

$$E = E_A + E_B. \tag{2.27}$$

According to quantum theory, the wavefunction is therefore *multiplicatively separable* and the energy *additively separable*.[79] As in the thermodynamics literature, a property that scales with the system size (such as the energy) is called *size-extensive*; properties that are independent of it are *size-intensive*.[80]

A unique and prominent feature of coupled cluster theory is that it seamlessly exhibits the proper scaling with system size. We prove that this is so for the energy and wavefunction. The recurring operator

$$\overline{H} = e^{-T} H e^{T} \tag{2.28}$$

is known as the *similarity transformed Hamiltonian*.[1] By the wavefunction's exponential form, multiplicative separability means that $T = T_A \otimes \mathbb{I} + \mathbb{I} \otimes T_B$. The identity $\overline{H} = \overline{H}_A \otimes \mathbb{I} + \mathbb{I} \otimes \overline{H}_B$ immediately follows. Performing the projection onto $|\mu_A \mu_B\rangle$, the amplitude equations are seen to be valid with an energy equal to $E = E_A + E_B$. Less concise proofs of the theory's size-extensivity have been presented in the literature.[1,81]

2.2.2. Molecular properties and excited states in coupled cluster theory

The extension of coupled cluster theory to properties and excited states began with the realization[18] that $\langle H \rangle$ (however it is defined) must be stationary for the Hellmann-Feynman theorem to remain valid. However, a proper generalization of the ground state theory requires that $\langle H \rangle$ be defined such that it equals the energy *and* satisfies the amplitude equations at its stationary points. By the prescripts of constrained optimization (see *e.g.* Adams[82]) we are led to define the coupled cluster expectation value as

$$\langle H \rangle_{\mathrm{cc}} = \langle \mathrm{R} | \overline{H} | \mathrm{R} \rangle + \sum_{\mu} \overline{t}_{\mu} \langle \mu | \overline{H} | \mathrm{R} \rangle = \frac{\langle \psi_{\Lambda} | H | \psi \rangle}{\langle \psi_{\Lambda} | \psi \rangle}, \tag{2.29}$$

where the *dual state*[19] is $\langle \psi_{\Lambda} | = \langle \mathrm{R} | + \sum_{\mu} \overline{t}_{\mu} \langle \mu | e^{-T}$ and $|\psi\rangle = e^{T} |\mathrm{R}\rangle$. The numbers $\overline{t}_{\mu}$ are called *multipliers* and are not to be confused with t_{μ}^{*}. At stationarity, we have the amplitude equations $\partial \langle H \rangle_{\mathrm{cc}} / \partial \overline{t}_{\mu} = 0$ and the *multiplier equations* $\partial \langle H \rangle_{\mathrm{cc}} / \partial t_{\mu} = 0$. The latter assumes the matrix form[21]

$$\boldsymbol{\eta}^{\dagger} + \overline{\mathbf{t}}^{\dagger} \mathbf{A} = 0, \tag{2.30}$$

where $A = \overline{H} - E$ is the *coupled cluster Jacobian operator*, $\mathbf{A}$ is the matrix representation of A in the basis $\{|\mu\rangle\}_{\mu}$, and $\eta_{\mu}^{\dagger} = \langle \mathrm{R} | \overline{H} | \mu \rangle$.

The correct definition of the expectation value paved the way for a sound coupled cluster response theory (CCRT). Since both the state and its dual has to be time-evolved, Koch and Jørgensen[19] considered the vector and functional forms of the Schrödinger equation:

$$H |\Psi\rangle = i \frac{\mathrm{d}}{\mathrm{d}t} |\Psi\rangle, \quad \langle \Psi_{\Lambda} | H = -i \frac{\mathrm{d}}{\mathrm{d}t} \langle \Psi_{\Lambda} |. \tag{2.31}$$

In time-independent theory, $|\psi\rangle$ and $\langle \psi_{\Lambda} |$ are equivalently determined by projecting $H |\psi\rangle = E |\psi\rangle$ and $\langle \psi_{\Lambda} | H = \langle \psi_{\Lambda} | E$ onto certain subspaces of $\mathscr{F}$.

We refer to these subspaces as the *dual space* $\mathscr{F}_\Lambda$ and the *state space* $\mathscr{F}_\psi$:

$$\mathscr{F}_\Lambda = \text{span}\,\{|\text{R}\rangle, |\mu\rangle\}_\mu, \tag{2.32}$$

$$\mathscr{F}_\psi = \text{span}\,\{e^T|\text{R}\rangle, e^T|\mu\rangle\}_\mu. \tag{2.33}$$

The determinants $|\mu\rangle$ are here restricted to excitation operators in the truncated $T = T_1 + \ldots + T_n$. Thus, $\mathscr{F}_\Lambda$ and $\mathscr{F}_n$ are the same; we assume n to be fixed and denote this space by $\mathscr{F}_\Lambda$ in the remainder of the text.

Consistently applying the same prescription in the time-dependent picture, the $|\Psi\rangle$ equation is projected onto $\mathscr{F}_\Lambda$ and the $\langle\Psi_\Lambda|$ equation onto $\mathscr{F}_\psi$. These projected equations are thereafter solved in orders of the perturbation:

$$t_\mu(t) = t_\mu^{(0)} + \sum_m t_\mu^{(m)}(t), \quad \bar{t}_\mu = \bar{t}_\mu^{(0)} + \sum_m \bar{t}_\mu^{(m)}(t). \tag{2.34}$$

The time evolved states are substituted in $\langle A\rangle_{\text{cc}} = \langle\Psi_\Lambda|A|\Psi\rangle$, from which the response functions are identified. For periodic V, the response formalism may also be derived by the quasi-energy Lagrangian approach.[21,57]

A different extension to excited states and properties is the equation of motion approach to coupled cluster theory (EOM-CC).[22] Its starting point is the observation that the true linear response function only depends on the excited states of the reference Hamiltonian H_0. An explicit construction of the excited states and their duals is performed by expanding them as

$$|\psi\rangle = e^T\Big(\sum_{\mu\geq 0} c_\mu|\mu\rangle\Big), \quad \langle\psi_\Lambda| = \Big(\sum_{\mu\geq 0}\bar{c}_\mu\langle\mu|\Big)e^{-T}, \tag{2.35}$$

for a fixed cluster operator T, *i.e.* by letting $|\psi\rangle \in \mathscr{F}_\psi$ and $|\psi_\Lambda\rangle \in \mathscr{F}_\Lambda$. In analogy with the ground state procedure, stationarity of the expectation value determines the excited states $|\psi\rangle$ and their duals $\langle\psi_\Lambda|$. The equations of motion $\partial\langle H\rangle_{\text{cc}}/\partial\bar{c}_\mu = 0$ and $\partial\langle H\rangle_{\text{cc}}/\partial c_\mu = 0$ assume the matrix forms[1]

$$\bar{\mathbf{c}}_n\,\mathbf{A} = \omega_n\,\bar{\mathbf{c}}_n, \quad \mathbf{A}\,\mathbf{c}_n = \omega_n\,\mathbf{c}_n, \tag{2.36}$$

where ω_n is the excitation energy of the wavefunction $|\psi_n\rangle$ whose coefficients are given by $\mathbf{c}_n$. Transition matrix elements are obtained by substituting the obtained states and duals in the exact expression:

$$\Gamma_{0\to k}^{V^{\omega_k}}\,\Gamma_{k\to 0}^A = \langle\psi_0^\Lambda|V^{\omega_k}|\psi_k\rangle\langle\psi_k^\Lambda|A|\psi_0\rangle. \tag{2.37}$$

More will be said on this topic in Chapter 4. We should note that while the excitation energies are identical in the two approaches,[19] their properties and transition elements differ.[64,65]

For a non-interacting composite AB system, action of H on $|\psi_k^A\rangle|\psi_0^B\rangle$ (we suppress the '$\otimes$') demonstrates that it is an excited state with excitation energy ω_k^A. The right transition element associated with it is, moreover,

$$\Gamma_{0\to k_A}^{\mathcal{O}} = \langle\psi_0^A|\langle\psi_0^B|(\mathcal{O}_A\otimes\mathbb{I} + \mathbb{I}\otimes\mathcal{O}_B)|\psi_k^A\rangle|\psi_0^B\rangle = \langle\psi_0^A|\mathcal{O}_A|\psi_k^A\rangle. \tag{2.38}$$

These are examples of size-intensive quantities; they do not scale with the size of the system. A similar derivation shows that the linear response functions are size-extensive. It is worth noting that equation of motion theory does not provide size-intensive transition elements[64] nor size-extensive response functions.[65] These facts count in favor of a time-dependent response theoretical approach, especially for large systems. The two theories give similar results on small systems.[21]

2.3. THE IMPORTANCE OF SYMMETRY

2.3.1. How symmetry is lost in coupled cluster theory and a survey of its detrimental consequences

It is widely known that the projection principle is responsible for coupled cluster theory's non-Hermiticity.[22,49,56,57] The most general expression for the energy illustrates that this is its origin:

$$E = \frac{\langle \varphi | H | \psi \rangle}{\langle \varphi | \psi \rangle}. \tag{2.39}$$

This relation is valid for the coupled cluster state $|\psi\rangle = e^T |R\rangle$ and any vector $|\varphi\rangle$ in $\mathscr{F}_\Lambda$ with $\langle \varphi | \psi \rangle \neq 0$. As the coupled cluster state $|\psi\rangle$ always contains determinants not included in $\mathscr{F}_\Lambda$ for truncated T it is not possible to write $|\varphi\rangle = |\psi\rangle$. There is consequently no guarantee that E is a real number.

The extensions to excited states and properties inherit the non-Hermiticity introduced by the projection principle. In fact, the unique expectation value consistent with the Hellmann-Feynman theorem and the amplitude equations is not necessarily real:

$$\langle H \rangle_{\text{cc}} = \langle \psi_\Lambda | H | \psi \rangle \neq \langle \psi | H | \psi_\Lambda \rangle = \langle H \rangle_{\text{cc}}^*. \tag{2.40}$$

Coupled cluster theory's excitation energies ω_n are moreover eigenvalues of the Jacobian $\mathbf{A}$, a non-Hermitian matrix: $\mathbf{A} \neq \mathbf{A}^\dagger$. Such a matrix can have complex eigenvalues and hence coupled cluster theory can predict complex excitation energies. Inconsistencies are unfortunately not restricted to the energies. Assuming real unperturbed eigenstates $\{|\psi_n\rangle\}_n$, in particular, it is not difficult to see that the exact relation

$$\langle\langle A; B \rangle\rangle_\omega = \langle\langle B; A \rangle\rangle_\omega, \quad A = A^\dagger, \quad B = B^\dagger, \tag{2.41}$$

is *not* satisfied in coupled cluster response theory. For instance, this non-physical artifact violates the symmetry of the frequency-dependent electric dipole polarizability $\alpha_{ij}(\omega) = \langle\langle \mu_i; \mu_j \rangle\rangle_\omega = \langle\langle \mu_j; \mu_i \rangle\rangle_\omega = \alpha_{ji}(\omega)$ and predicts, incorrectly, that the velocity gauge rotary strength is origin dependent.[57]

As the Hamiltonian operator H is Hermitian, it has an orthonormal basis of eigenstates $|\psi_n\rangle$ for the space on which it acts (see Theorem 6.25 of

Friedberg *et al.*[71]). The projectors $|\psi_n\rangle\langle\psi_n|$ onto the associated eigenspaces are instrumental in forming the spectral representation of $\langle\langle A; B\rangle\rangle_\omega$ as they justify the relation $\mathbb{I} = \sum_n |\psi_n\rangle\langle\psi_n|$. These true projectors are however substituted by the 'pseudo-projectors' $|\psi_n\rangle\langle\psi_n^\Lambda|$ in equation of motion theory. This inconsistency vanishes in our symmetric formulation (see Chapter 4).

Also the loss of variationality may be understood as rooted in the loss of Hermiticity by projection: the expectation value $\langle\psi|H|\psi\rangle$ always exceeds the ground state energy, but this is emphatically not the case for $\langle\psi_\Lambda|H|\psi\rangle$.[58] Note that the dual $|\psi_\Lambda\rangle$ was necessary to keep the amplitude equations valid in the definition of the expectation value. It should be stressed that symmetry does not guarantee a variational energy, however, as the variational theorem pertains only to expectation values of Hermitian operators. The ground state theory presented in this thesis is symmetric but not variational.

2.3.2. Conical intersections, excited state dynamics, and the proper description of photochemical phenomena

An analysis of the quantum state $\Psi(\mathbf{r}, \mathbf{R})$ of molecular systems begins with the full time-dependent Schrödinger equation. Let us partition H as

$$H(\mathbf{r}, \mathbf{R}) = T_n(\mathbf{R}) + T_e(\mathbf{r}) + U(\mathbf{r}, \mathbf{R}) = T_n(\mathbf{R}) + H_e(\mathbf{r}, \mathbf{R}), \qquad (2.42)$$

where T_n and T_e denote the nuclear and electronic kinetic energy operators, $\mathbf{r}$ and $\mathbf{R}$ are the collection of electronic and nuclear coordinates, and U contains all potential terms involving the nuclei and electrons.

The Hilbert space $\mathscr{H}$ is the space of square integrable functions over the electronic and nuclear coordinates. It may be written $\mathscr{H} = \mathscr{H}(\mathbf{r}) \otimes \mathscr{H}(\mathbf{R})$, where $\mathscr{H}(\mathbf{r})$ and $\mathscr{H}(\mathbf{R})$ are referred to as the *electronic* and *nuclear spaces*.[83] For fixed nuclear geometries $\mathbf{R}$,

$$H_e(\mathbf{r}, \mathbf{R})\, \Phi_i(\mathbf{r}, \mathbf{R}) = E_i(\mathbf{R})\, \Phi_i(\mathbf{r}, \mathbf{R}) \qquad (2.43)$$

is solved for the *electronic states* $\Phi_i(\mathbf{r}, \mathbf{R})$, an orthonormal basis for the electronic space: $\int \mathrm{d}\mathbf{r}\, \Phi_i(\mathbf{r}, \mathbf{R})^*\Phi_j(\mathbf{r}, \mathbf{R}) = \langle i(\mathbf{R})|j(\mathbf{R})\rangle = \delta_{ij}$. The state $\Psi(\mathbf{r}, \mathbf{R})$ thus has the useful representation

$$\Psi(\mathbf{r}, \mathbf{R}) = \sum_i \Phi_i(\mathbf{r}, \mathbf{R})\, \chi_i(\mathbf{R}), \qquad (2.44)$$

where the χ_i are called *nuclear wavefunctions*.[35] Upon multiplication by $\Phi_j(\mathbf{r}, \mathbf{R})^*$ and integration over $\mathbf{r}$ the Schrödinger equation may be written

$$\big(T_n(\mathbf{R}) + E_j(\mathbf{R})\big)\chi_j(\mathbf{R}) - \sum_i \Lambda_{ji}(\mathbf{R})\chi_i(\mathbf{R}) = i\frac{\mathrm{d}\chi_j(\mathbf{R})}{\mathrm{d}t}, \qquad (2.45)$$

where the *non-adiabatic couplings* $\Lambda_{ji}(\mathbf{R}) = -\delta_{ji}\, T_n(\mathbf{R}) + \langle j(\mathbf{R})|T_n(\mathbf{R})\, i(\mathbf{R})\rangle$ are operators on nuclear space. They are responsible for the system's coupling of nuclear and electronic motion.

The nature of the non-adiabatic couplings is best understood in terms of the *derivative* and *scalar couplings*, defined as $\mathbf{F}_{ji} = \langle j(\mathbf{R})|\boldsymbol{\nabla}i(\mathbf{R})\rangle$ and $G_{ji} = \langle j(\mathbf{R})|\boldsymbol{\nabla}^2 i(\mathbf{R})\rangle$. In fact[83]

$$\Lambda_{ji}(\mathbf{R}) = \frac{1}{2M}\Big(\mathbf{F}_{ji}\cdot\boldsymbol{\nabla} + G_{ji}\Big), \tag{2.46}$$

where M is an averaged nuclear mass. It is interesting that Equation (2.45) may be written as[84]

$$\Big(-\frac{1}{2M}(\boldsymbol{\nabla} + \mathbf{F})^2 + \mathbf{E}\Big)\boldsymbol{\chi} = i\frac{\mathrm{d}\boldsymbol{\chi}}{\mathrm{d}t}, \tag{2.47}$$

where the χ_i have been collected in the vector $\boldsymbol{\chi}$, $\mathbf{F}$ is the matrix of derivative couplings $\mathbf{F}_{ij}$, and $\mathbf{E}$ is a diagonal matrix of the electronic energies E_i.

This equation provides a natural starting point at which to introduce approximations. The *group Born-Oppenheimer approximation*[83] truncates the representation as $\Psi(\mathbf{r}, \mathbf{R}) = \sum_{i\in g}\Phi_i(\mathbf{r},\mathbf{R})\chi_i(\mathbf{R})$, where g is a group of electronic states. The Schrödinger equation then assumes the form

$$\Big(-\frac{1}{2M}(\boldsymbol{\nabla} + \mathbf{F}^{(g)})^2 + \mathbf{W}^{(g)}\Big)\boldsymbol{\chi}^{(g)} = i\frac{\mathrm{d}\boldsymbol{\chi}^{(g)}}{\mathrm{d}t}, \tag{2.48}$$

where the superscript (g) denotes the restriction of the matrices and vectors to the group g, and

$$W_{ji}^{(g)}(\mathbf{R}) = E_i^{(g)}(\mathbf{R})\,\delta_{ij} + \frac{1}{2M}\sum_{k\notin g}(\mathbf{F}_{kj})^*\cdot\mathbf{F}_{ki} \tag{2.49}$$

is the *dressed potential matrix*. The *group adiabatic approximation*[35] ignores the non-group coupling in $\mathbf{W}$ (the sum over $k\notin g$) such that $\mathbf{W}^{(g)} = \mathbf{E}^{(g)}$. The *adiabatic approximation* restricts the group g to a single electronic state.

Of particular interest are the adiabatic approximations, as the group Born-Oppenheimer expansion $\Psi(\mathbf{r},\mathbf{R}) = \sum_{i\in g}\Phi_i(\mathbf{r},\mathbf{R})\chi_i(\mathbf{R})$ is only tenable when the adiabatic assumption is valid.[35] Since $\mathbf{F}_{ij}$ may be written

$$\mathbf{F}_{ij} = \frac{\langle i(\mathbf{R})|\,(\boldsymbol{\nabla}H_e(\mathbf{r},\mathbf{R}))\,j(\mathbf{R})\rangle}{E_i(\mathbf{R}) - E_j(\mathbf{R})} \tag{2.50}$$

the group adiabatic approximation is not tenable when an out-group state Φ_k becomes electronically degenerate with an in-group state Φ_i (the $\mathbf{F}_{ki}$ can no longer be ignored). In the following we therefore assume that all relevant electronic states are included in the group such that they may be treated by the group adiabatic approximation $\mathbf{W}^{(g)} = \mathbf{E}^{(g)}$. The superscript (g) will moreover be suppressed.

If we restrict ourselves to some region of nuclear geometries $\mathbf{R}$, and include all relevant electronic states Φ_i, $i\in g$, the nuclear dynamics may in

principle be understood from the potential energy surfaces $\mathbf{E}$ and the derivative couplings $\mathbf{F}$. These quantities can be estimated by quantum chemical methods, *e.g.* by coupled cluster theory. In organic molecules, conical intersections between electronic states are ubiquitous.[35] An *ab initio* approach to nuclear dynamics is therefore absolutely essential to excited state dynamics. We refer the reader to Ben-Nun *et al.*[45] for more on the use of quantum chemistry in non-adiabatic regimes.

Let us suppose that the energies $\mathbf{E}$ are determined as the eigenvalues of an arbitrary matrix $\mathbf{H}$. If only two states couple strongly, we may restrict our attention to the reduced 2×2 problem, here written in the Hermitian case $\mathbf{H} = \mathbf{H}^\dagger$:

$$\mathbf{H} = \begin{pmatrix} \overline{E} - \Delta & S \\ S^* & \overline{E} + \Delta \end{pmatrix}. \tag{2.51}$$

The eigenvalues are $E_\pm = \overline{E} \pm \sqrt{\Delta^2 + |S|^2}$. Evidently, the two energies become degenerate if and only if $\Delta = 0$ and $S = 0$. The nuclear coordinate $\mathbf{R}$ may inhabit either the *branching plane* $\mathcal{M}$ or the *intersection space* $\mathcal{I}$:

$$\mathcal{M} = \{\mathbf{R} : \Delta^2 + S^2 > 0\}, \tag{2.52}$$
$$\mathcal{I} = \{\mathbf{R} : \Delta = 0 \text{ and } S = 0\}. \tag{2.53}$$

In $\mathcal{I}$, two independent requirements are set, implying that $\dim(\mathcal{I}) = F - 2$ for a system with F internal nuclear degrees of freedom. Consequently, the branching space is two-dimensional, $\dim(\mathcal{M}) = 2$. Due to the double-cone shape of $E_\pm = E_\pm(\Delta, S)$ the degeneracy is named a *conical intersection*.[35]

In coupled cluster theory, the energies $\mathbf{E}$ are derived from the eigenvalue problem associated with a non-Hermitian matrix. Let us therefore consider an arbitrary 2×2 matrix $\mathbf{H}$:

$$\mathbf{H} = \begin{pmatrix} \overline{E} - \Delta & S - A \\ S + A & \overline{E} + \Delta \end{pmatrix}. \tag{2.54}$$

In this case the eigenvalues are $E_\pm = \overline{E} \pm \sqrt{\Delta^2 + S^2 - A^2}$. Except perhaps at the degeneracy, $\mathbf{H}$ is diagonalizable because it has distinct eigenvalues (see Theorem 5.2.3 in Anton and Rorres[85]). There thus exists an invertible matrix $\mathbf{M}$ such that $\mathbf{M}\,\mathbf{H}\,\mathbf{M}^{-1}$ is diagonal with elements $\overline{E} \pm \epsilon$. Transforming back to the original matrix, we find that if

$$\mathbf{M} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \tag{2.55}$$

then $\mathbf{H}$ may be written as in Equation (2.54) with

$$\Delta = \eta^{-1}(ad + bc)\,\epsilon, \tag{2.56}$$
$$S = \eta^{-1}(ca - bd)\,\epsilon, \tag{2.57}$$
$$A = \eta^{-1}(ca + bd)\,\epsilon, \tag{2.58}$$

where $\eta = \det(\mathbf{M})$. *If* the matrix $\mathbf{H}$ remains diagonalizable at the degeneracy $\epsilon = 0$, we evidently must have $\Delta = 0$, $S = 0$, and $A = 0$.

In non-Hermitian theories we thus have three relevant spaces for the nuclear coordinates, as they inhabit either the *real* $\mathcal{M}_\mathbb{R}$ or *complex* $\mathcal{M}_\mathbb{C}$ branching spaces, or the intersection space $\mathcal{I}$:

$$\mathcal{M}_\mathbb{R} = \{\mathbf{R} : \Delta^2 + S^2 > A^2\}, \tag{2.59}$$

$$\mathcal{M}_\mathbb{C} = \{\mathbf{R} : \Delta^2 + S^2 < A^2 \text{ or } \Delta^2 + S^2 - A^2 \in \mathbb{C}\}, \tag{2.60}$$

$$\mathcal{I} = \{\mathbf{R} : \Delta = 0, S = 0, \text{ and } A = 0\}. \tag{2.61}$$

Hättig[48] and Köhn and Tajti[49] have noted $\mathcal{I}$'s three conditions, concluding that non-Hermitian theories predict, incorrectly, that $\dim(\mathcal{I}) = F - 3$ for conical intersections $\mathcal{I}$. To date, we are not aware of any definite numerical evidence nor an analysis of the independence of Δ, S, and A. We note that this independence must break down in the full expansion $T = T_1 + T_2 + \ldots + T_N$, as the eigenvalues of the non-Hermitian $\mathbf{A}$ and the Hermitian H merge in this limit.

Of especial concern is the possibility of complex energies near an intersection. Most often $\mathcal{M}_\mathbb{R}$ is encountered in coupled cluster calculations, and the $\mathbf{H}$ block produces two real energies, as it should. In the vicinity of an intersection $\mathcal{I}$, where $\Delta^2 + S^2 - A^2 \approx 0$, we might encounter $\mathbf{R} \in \mathcal{M}_\mathbb{C}$ if this quantity dips below zero. This has in fact been observed for formaldehyde, where a complex pair is formed and subsequently vanishes close to the $2\,^1\mathrm{A}_1/3\,^1\mathrm{A}_1$ intersection.[49]

A few remarks on the question of diagonalizability at $\epsilon = 0$ are in order. In equation of motion theory, the states are substituted in $\langle\langle A; V^\omega \rangle\rangle_\omega$. As this recipe is based on the orthonormality and completeness of the eigenstates, an occurence of parallell eigenvectors (*i.e.* non-diagonalizability[85]) would likely place us outside its domain of validity. In response theory, on the other hand, non-diagonalizability results in poles of higher order in the linear response function, a pathological behavior best remedied by including more excitations in T.[14] We therefore deem of little relevance in practice the case $\Delta^2 + S^2 = A^2 \neq 0$ considered by Hättig[48] as it cannot be obtained with the assumption that $\mathbf{H}$ is diagonalizable at the degeneracy.

2.4. LIMITATIONS AND ALTERNATIVES

2.4.1. The limitations of coupled cluster theory

Whenever the true wavefunction contains several leading determinants we have entered the multireference regime, where the accuracy of models that favor a particular determinant quickly deteriorates. Let us explain the origins of the breakdown of coupled cluster theory for the singles and doubles truncation of T. We may write $|\psi\rangle = e^T |\mathrm{R}\rangle = \sum_{\mu \geq 0} c_\mu |\mu\rangle$ with

$$c_0 = 1, \quad c_i^a = t_i^a, \quad c_{ij}^{ab} = t_{ij}^{ab} + t_i^a t_j^b, \quad \ldots \tag{2.62}$$

The triply and higher excited coefficients arise from disconnected contributions in e^T, e.g. $T_1 T_2$ and $\frac{1}{6} T_1^3$ in the case of triply excited determinants. When the reference ceases to be the dominant term the cluster amplitudes t_μ often grow and the disconnected contributions can, by their increasing importance, lead to unpredictable results.[31] This unpredictability affects the properties and transition moments to a higher extent, as the energy only depends on T_2 and T_1^2 by construction: $E = \langle R \,|\, H e^T \,|\, R \rangle$.

The inadequacy of the reference determinant may to some extent be alleviated by simultaneously rotating its spin orbitals, as is done in the orbital-optimized[86,87] and Brueckner[88] coupled cluster theories. It should however be understood that an accurate single-reference description requires that the $1/r_{12}$ perturbational contribution can be made 'small' by a proper choice of the reference.[31] It can therefore only provide meaningful results when the description of correlation is well approximated by a mean field or independent particle approximation. For multireference coupled cluster approaches, the reader is referred to the review by Lyakh et al.[31]

2.4.2. Alternative symmetric coupled cluster theories: the variational and quasi-variational approaches

One of the major obstacles in the formulation of symmetric coupled cluster theories is the prohibitive scaling with the number of electrons. A key point is that the T^4 terminating similarity transformed Hamiltonian of the traditional formulation is not reproduced:[56]

$$ e^{-T} H e^T = H + [H, T] + \frac{1}{2}[[H, T], T] + \ldots + \frac{1}{24}[[[[H, T], T], T], T]. \quad (2.63) $$

These theories instead encounter analogous never-ending series, necessitating truncation schemes for practical purposes.

A first and obvious approach is the variational one, i.e. using the symmetric expectation value $\langle H \rangle$ for the coupled cluster state, fixing its amplitudes t_μ by enforcing stationarity. However,

$$ \langle H \rangle = \frac{\langle R \,|\, e^{T^\dagger} H e^T \,|\, R \rangle}{\langle R \,|\, e^{T^\dagger} e^T \,|\, R \rangle} \quad (2.64) $$

is an infinite series in T and $T^\dagger$.[61] Quite a few truncation schemes have been introduced in the literature, where we mention those of Bartlett and Noga[89] and Robinson and Knowles.[90] These models are best called quasi-variational, as they approximate $\langle H \rangle$ and are hence no longer strictly variational.

A formulation more reminiscent of traditional coupled cluster theory is the unitary approach of Yaris.[63] As above, the exact symmetric expectation values is considered, but the exponential ansatz is modified:

$$ |\psi\rangle = e^\sigma \,|R\rangle, \quad \sigma^\dagger = -\sigma. \quad (2.65) $$

If linearity in τ_μ and $\tau_\mu^\dagger$ is assumed, the unique choice is the commonly adopted[62] $\sigma = T - T^\dagger$. Regardless of σ, the exact expectation value assumes the familiar and attractive form

$$\langle H \rangle = \langle \mathrm{R} | e^{-\sigma} H e^{\sigma} | \mathrm{R} \rangle. \tag{2.66}$$

The commutator series never ends, however,

$$e^{-\sigma} H e^{\sigma} = H + [H, \sigma] + \frac{1}{2}[[H, \sigma], \sigma] + \ldots \tag{2.67}$$

Of note are the proposed truncation schemes set forth by Bartlett and coworkers[91] and by Kutzelnigg.[92] For an excellent review of the alternative coupled cluster formulations (both Hermitian and non-Hermitian) the reader may consult Szalay *et al.*[56]

Chapter 3

Symmetric coupled cluster theory

3.1. THEORY

3.1.1. The dual and state spaces revisited

The accuracy of the two-fold state-dual description is expected to deteriorate whenever the duals are poor representations of the states. This is especially evident in equation of motion theory, where the transition elements are in effect obtained by replacing $\langle \psi_0 |$ by $\langle \psi_0^\Lambda |$:

$$\Gamma_{0 \to n}^A = \langle \psi_0^\Lambda | A | \psi_n \rangle. \tag{3.1}$$

We will advocate an approach which is best understood from a geometrical perspective. First recall that the ground and excited states $|\psi\rangle$ and their duals $|\psi_\Lambda\rangle$ inhabit the state and dual spaces, $\mathscr{F}_\psi$ and $\mathscr{F}_\Lambda$:

$$\mathscr{F}_\psi = \mathrm{span}\,\{e^T |\mathrm{R}\rangle, e^T |\mu\rangle\}_\mu, \quad \mathscr{F}_\Lambda = \mathrm{span}\,\{|\mathrm{R}\rangle, |\mu\rangle\}_\mu. \tag{3.2}$$

We refer to the listed sets as the *standard* bases of $\mathscr{F}_\psi$ and $\mathscr{F}_\Lambda$. A theorem states that the best approximation of $|\psi\rangle$ in the dual space $\mathscr{F}_\Lambda$ is the orthogonal projection $P|\psi\rangle$ onto $\mathscr{F}_\Lambda$, as measured by the norm of the error.[67] We give a proof in Appendix A. Thus, $|\psi_\Lambda\rangle$ is not the optimal approximation of $|\psi\rangle$ if it differs from $P|\psi\rangle$. Figure 3.1 illustrates the relevant objects.
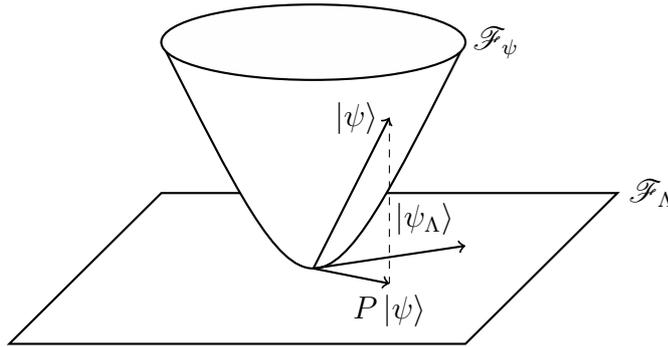
FIGURE 3.1: *An illustration of the dual space $\mathscr{F}_\Lambda$ and state space $\mathscr{F}_\psi$.*

3.1.2. Two approaches to the expectation value

Expectation values may be introduced by determining a dual state $\langle \psi_\Lambda |$ consistent with an existing state $|\psi\rangle$. This is the approach taken in coupled

cluster theory, where the dual is introduced *a posteriori* for a state already fixed by the projection principle. Alternatively, the wavefunction $|\psi\rangle$ can be determined from stationarity of $\langle H \rangle$ or an approximation of it, as is done in the quasi-variational coupled cluster formulations. This is the approach we employ in this thesis.

Let us explain the motivation for these approaches. The coupled cluster expectation value is defined as $\langle H \rangle_{\mathrm{cc}} = \langle \psi_\Lambda | H | \psi \rangle$, where $H = H_0 + \lambda V$. The state $|\psi\rangle$ and its dual $\langle \psi_\Lambda |$ are determined by stationarity of $\langle H \rangle_{\mathrm{cc}}$ with respect to t_μ and $\bar{t}_\mu$ at $\lambda = 0$. Now note that

$$\frac{\mathrm{d}\langle H \rangle_{\mathrm{cc}}}{\mathrm{d}\lambda} = \langle V \rangle_{\mathrm{cc}} + \sum_\mu \Big(\frac{\mathrm{d}t_\mu}{\mathrm{d}\lambda} \frac{\partial \langle H \rangle_{\mathrm{cc}}}{\partial t_\mu} + \frac{\mathrm{d}\bar{t}_\mu}{\mathrm{d}\lambda} \frac{\partial \langle H \rangle_{\mathrm{cc}}}{\partial \bar{t}_\mu} \Big) \tag{3.3}$$

by an application of the elementary chain rule. This expectation value definition is consistent with the Hellmann-Feynman theorem because every contribution in the sum vanishes at $\lambda = 0$ by stationarity; they vanish by the amplitude and multiplier equations (see Section 2.2.2).

A different result is obtained when the expectation value is understood to be the exact $\langle H \rangle = \langle \psi | H | \psi \rangle / \langle \psi | \psi \rangle$. By differentiation we find the similar

$$\frac{\mathrm{d}\langle H \rangle}{\mathrm{d}\lambda} = \langle V \rangle + \sum_\mu \Big(\frac{\mathrm{d}t_\mu}{\mathrm{d}\lambda} \frac{\partial \langle H \rangle}{\partial t_\mu} + \frac{\mathrm{d}t_\mu^*}{\mathrm{d}\lambda} \frac{\partial \langle H \rangle}{\partial t_\mu^*} \Big). \tag{3.4}$$

However, as $\partial \langle H \rangle / \partial t_\mu^* = (\partial \langle H \rangle / \partial t_\mu)^*$, $\partial \langle H \rangle / \partial t_\mu |_{\lambda=0} = 0$ is a sufficient condition to obtain the Hellmann-Feynman theorem. The above admits a generalization to a class $\langle H \rangle_{F,G}$ of expectation value approximations in agreement with the Hellmann-Feynman theorem,

$$\langle H \rangle_{F,G} = \frac{\langle \psi | F(H) | \psi \rangle}{\langle \psi | G | \psi \rangle}, \quad \frac{\partial \langle H \rangle_{F,G}}{\partial t_\mu} \Big|_{\lambda=0} = 0, \tag{3.5}$$

where G is Hermitian and F is a Hermitian operator-valued function of H that satisfies $\partial F(H)/\partial \lambda |_{\lambda=0} = F(V)$. A proof is given in Appendix A. The theory introduced in this thesis represents one particular choice among all the possible definitions $\langle H \rangle_{F,G}$.

3.1.3. The symmetric coupled cluster *ansatz*

The symmetric coupled cluster (SCC) *ansatz* is the two-fold requirement that the dual state $|\psi_\Lambda\rangle$ is the optimal approximation in $\mathscr{F}_\Lambda$ of the exponential coupled cluster state $|\psi\rangle$:

$$|\psi\rangle = e^T |\mathrm{R}\rangle, \quad |\psi_\Lambda\rangle = P |\psi\rangle. \tag{3.6}$$

To determine one state (t_μ) instead of two $(t_\mu, \bar{t}_\mu)$ we define an expectation value in the class $\langle H \rangle_{F,G}$. This we obtain by letting the dual state play both

the role of ket ($|\psi_\Lambda\rangle$) and bra vector ($\langle\psi_\Lambda|$), averaging the two:

$$\langle H\rangle_{\mathrm{scc}} = \Re e\frac{\langle\psi_\Lambda|H|\psi\rangle}{\langle\psi_\Lambda|\psi\rangle} = \frac{\langle\psi|\frac{1}{2}(PH+HP)|\psi\rangle}{\langle\psi|P|\psi\rangle}. \tag{3.7}$$

Truncating the cluster operator as $T = T_1 + T_2 + \ldots + T_n$ the amplitudes are determined from the *symmetric amplitude equations*

$$\frac{\partial\langle H\rangle_{\mathrm{scc}}}{\partial t_\mu^*} = 0. \tag{3.8}$$

The *symmetric coupled cluster energy* E_{scc} is the value of $\langle H\rangle_{\mathrm{scc}}$ for the amplitudes which solve $\partial\langle H\rangle_{\mathrm{scc}}/\partial t_\mu^* = 0$. By truncation of T, a hierarchy of models are defined (SCCS, SCCSD, SCCSDT, ...) that converges to an exact solution in $\mathscr{F}$ when $T = T_1 + T_2 + \ldots + T_N$. For ease of discussion we will frequently refer to the above as the 'symmetric theory,' reserving the term 'traditional' for the projective approach described in Section 2.2.

Let us prove that symmetric coupled cluster theory is exact in the limit of an untruncated T. Note that $P = \mathbb{I}$ by $\mathscr{F}_\Lambda = \mathscr{F}$, giving $\langle H\rangle_{\mathrm{scc}} = \langle H\rangle$. The requirement that $\langle H\rangle$ is stationary may then shown to be equivalent to the projection of $H|\psi\rangle = E|\psi\rangle$ onto the basis $\{e^T|\mathrm{R}\rangle, e^T|\mu\rangle\}_\mu$ of $\mathscr{F}$. The obtained $|\psi\rangle$ therefore satisfies $H|\psi\rangle = E|\psi\rangle$. This finishes the proof.

3.1.4. The energy and amplitude equations

The present section is devoted to the derivation of the matrix form of the energy and amplitude equations of the symmetric formulation. In more precise terms, we derive and present $\langle H\rangle_{\mathrm{scc}}$ and $\partial\langle H\rangle_{\mathrm{scc}}/\partial t_\mu^* = 0$ written in terms of vectors and matrices. These expressions are necessary for purposes of implementation and facilitate the comparison of the symmetric and traditional theories.

Let the standard basis $\{|\mathrm{R}\rangle, |\mu\rangle\}_\mu$ of $\mathscr{F}_\Lambda$ be orthonormal. The general case is postponed to Section 3.2.2. In this basis the orthogonal projection operator onto $\mathscr{F}_\Lambda$ assumes the form $P = \sum_{\mu\geq 0}|\mu\rangle\langle\mu|$.[*] The denominator of $\langle H\rangle_{\mathrm{scc}}$ may thus be written

$$\langle\psi|P|\psi\rangle = 1 + \sum_\mu\langle\mathrm{R}|e^{T^\dagger}|\mu\rangle\langle\mu|e^T|\mathrm{R}\rangle = 1 + \mathbf{q}^\dagger\mathbf{q}, \tag{3.9}$$

where we have let $q_\mu = \langle\mu|e^T|\mathrm{R}\rangle$.

A suitable expression for the numerator is obtained by introducing the similarity transformed Hamiltonian $\overline{H} = e^{-T}He^T$. This is done by noting

---

[*]This is verified by expanding any $|\varphi\rangle \in \mathscr{F}$ in an enlarged standard basis $\{|\mathrm{R}\rangle, |\mu\rangle, \ldots\}_\mu$. The vectors '...' may be selected from the orthogonal complement $\mathscr{F}_\Lambda^\perp$ as $\mathscr{F} = \mathscr{F}_\Lambda \oplus \mathscr{F}_\Lambda^\perp$. See Thorem 3.3-4 in Kreyszig.[75]

that we can write $P\,|\psi\rangle = e^{-T^\dagger} P\, e^{T^\dagger} P\,|\psi\rangle$ because $\mathscr{F}_\Lambda$ is closed under $T^\dagger$ and $P|_{\mathscr{F}_\Lambda} = \mathbb{I}$. As a consequence,

$$
\begin{aligned}
\langle\psi|\,H\,P\,|\psi\rangle &= \sum_{\nu\sigma\geq 0} \langle\mathrm{R}|\,\overline{H}^\dagger\,|\nu\rangle\langle\nu|\,e^{T^\dagger}\,|\sigma\rangle\langle\sigma|\,e^T\,|\mathrm{R}\rangle \\
&= (1 + \mathbf{q}^\dagger\mathbf{q})\,\overline{H}_{00}^\dagger + \boldsymbol{\xi}^\dagger\mathbf{Q}^\dagger\mathbf{q},
\end{aligned}
\tag{3.10}
$$

where we have defined $Q_{\mu\nu} = \langle\mu|\,e^T\,|\nu\rangle$ and $\xi_\mu = \langle\mu|\,\overline{H}\,|\mathrm{R}\rangle$. It is worthwhile to note that $\xi_\mu = 0$ are the traditional amplitude equations. We now take the real part of $\langle\psi|\,H\,P\,|\psi\rangle$ and divide it by $\langle\psi|\,P\,|\psi\rangle$ to obtain the symmetric coupled cluster expectation value

$$
\langle H\rangle_{\mathrm{scc}} = \Re e\,(\overline{H}_{00} + \overline{\boldsymbol{\tau}}^\dagger\boldsymbol{\xi}),
\tag{3.11}
$$

where $\overline{\boldsymbol{\tau}} = \mathbf{Q}^\dagger\mathbf{q}/(1 + \mathbf{q}^\dagger\mathbf{q})$ is an analogue of the multiplier vector $\overline{\mathbf{t}}$. We return to explain in what sense this vector is the natural analogue of $\overline{\mathbf{t}}$ in the next section. Using the amplitudes that satisfy $\partial\langle H\rangle_{\mathrm{scc}}/\partial t_\mu^* = 0$ the symmetric coupled cluster energy can be written $E_{\mathrm{scc}} = \Re e\,(\overline{H}_{00} + \overline{\boldsymbol{\tau}}^\dagger\boldsymbol{\xi})$.

We now turn to the symmetric amplitude equations. In order to evaluate $\partial\langle H\rangle_{\mathrm{scc}}/\partial t_\mu^* = 0$, it is most convenient to first derive the partial derivatives of the individual $\langle\psi|\,P\,|\psi\rangle$, $\langle\psi|\,HP\,|\psi\rangle$, and $\langle\psi|\,PH\,|\psi\rangle$ terms. By noting that $\partial|\psi\rangle/\partial t_\mu = e^T\,|\mu\rangle$, the definitions $N = 1 + \mathbf{q}^\dagger\mathbf{q}$ and $\eta_\mu^\dagger = \langle\mathrm{R}|\,\overline{H}\,|\mu\rangle$ provide us with the following expressions:

$$
\begin{aligned}
\frac{\partial}{\partial t_\mu^*}\langle\psi|\,P\,|\psi\rangle &= \sum_\nu \langle\mu|\,e^{T^\dagger}\,|\nu\rangle\langle\nu|\,e^T\,|\mathrm{R}\rangle \\
&= (\mathbf{Q}^\dagger\mathbf{q})_\mu,
\end{aligned}
\tag{3.12}
$$

$$
\begin{aligned}
\frac{\partial}{\partial t_\mu^*}\langle\psi|\,HP\,|\psi\rangle &= \sum_{\nu\sigma\geq 0}\langle\mu|\,\overline{H}^\dagger\,|\nu\rangle\langle\nu|\,e^{T^\dagger}\,|\sigma\rangle\langle\sigma|\,e^T\,|\mathrm{R}\rangle \\
&= (N\,\boldsymbol{\eta} + \overline{\mathbf{H}}^\dagger\mathbf{Q}^\dagger\mathbf{q})_\mu,
\end{aligned}
\tag{3.13}
$$

$$
\begin{aligned}
\frac{\partial}{\partial t_\mu^*}\langle\psi|\,PH\,|\psi\rangle &= \sum_{\nu\sigma\geq 0}\langle\mu|\,e^{T^\dagger}\,|\nu\rangle\langle\nu|\,e^T\,|\sigma\rangle\langle\sigma|\,\overline{H}\,|\mathrm{R}\rangle \\
&= (\mathbf{Q}^\dagger\mathbf{q}\,\overline{H}_{00} + \mathbf{Q}^\dagger\mathbf{Q}\,\boldsymbol{\xi})_\mu.
\end{aligned}
\tag{3.14}
$$

These are all the ingredients necessary to write the amplitude equations in familiar notation. Let us write them in a more convenient form:

$$
0 = \frac{\partial\langle H\rangle_{\mathrm{scc}}}{\partial t_\mu^*} = \frac{1}{N}\Big(\frac{1}{2}\frac{\partial}{\partial t_\mu^*}\langle\psi|\,PH + HP\,|\psi\rangle - E_{\mathrm{scc}}\,\langle\psi|\,P\,|\psi\rangle\Big).
\tag{3.15}
$$

We define the *symmetric* coupled cluster Jacobian as $A = \overline{H} - E_{\mathrm{scc}}$. By use of the derived partial derivatives of $\langle\psi|\,P\,|\psi\rangle$, $\langle\psi|\,HP\,|\psi\rangle$, and $\langle\psi|\,PH\,|\psi\rangle$, the symmetric amplitude equations can be shown to read

$$
\mathbf{0} = \boldsymbol{\eta} + \mathbf{A}^\dagger\overline{\boldsymbol{\tau}} + \alpha\,\overline{\boldsymbol{\tau}} + \frac{1}{N}\,\mathbf{Q}^\dagger\mathbf{Q}\,\boldsymbol{\xi},
\tag{3.16}
$$

where $\mathbf{A}$ is the $\{|\mu\rangle\}_\mu$-representation of A and $\alpha = \langle R|A|R\rangle$. We employ the same notation to describe the objects $(\alpha, \boldsymbol{\xi}, \boldsymbol{\eta}, \mathbf{A})$ associated with A in both the symmetric and traditional formalisms, even though they are different; the amplitudes entering $\overline{H}$ are different and the energy subtracted from it to form A differs. This should be kept in mind in the discussions that follow.

3.1.5. A comparison of the symmetric and traditional theories

Let us first understand what happens if we attempt to use the traditional amplitudes, derived by projection, in the energy and amplitude equations of the symmetric theory. The traditional amplitude equations imply that $\boldsymbol{\xi} = 0$, giving $E_{\text{scc}} = \Re e\,\overline{H}_{00}$. Thus, *if* the traditional energy $E_{\text{cc}} = \overline{H}_{00}$ is a real number, the two formulations agree on its value. But then α vanishes, allowing us to write the symmetric amplitude equations as

$$\mathbf{0} = \boldsymbol{\eta}^\dagger + \overline{\boldsymbol{\tau}}^\dagger \mathbf{A}. \tag{3.17}$$

The traditional and symmetric theories will thus agree if the traditional multiplier equation $\boldsymbol{\eta}^\dagger + \overline{\mathbf{t}}^\dagger \mathbf{A} = \mathbf{0}$ is satisfied with $\overline{\mathbf{t}} = \overline{\boldsymbol{\tau}}$. However, it turns out that this is most often not the case. The two theories thus provide different amplitudes and different energies in general.

A notable difference between the two formalisms is found in how they approximate the matrix elements of the Hamiltonian H. It is most convenient to see this by setting $\langle\varphi| = \langle\psi|P$ in Equation (2.39):

$$E_{\text{cc}} = \frac{\langle\psi|PH|\psi\rangle}{\langle\psi|P|\psi\rangle}. \tag{3.18}$$

Here $|\psi\rangle = e^T|R\rangle$ contains the traditional amplitudes, not the symmetric ones. For comparison, see the expression for the symmetric energy in Equation (3.7). We should mention that since $E_{\text{cc}} = \langle R|He^T|R\rangle$, no higher than doubly excited determinants contribute to its value. Although terms involving triples and quadruples do enter in Equation (3.18), they all cancel out due to the traditional amplitude equations.

At the singles and doubles truncation of T, the coupled cluster state $|\psi\rangle$ may be expanded in the reference, singly (S), doubly (D), triply (T), and quadrupluply (Q) excited determinants (quintuples and higher are destroyed by P and do not contribute). Evaluating the energy term by term we see that the inclusion of triples and quadruples differs in the two formulations:

$$\text{CC}: \begin{cases} \langle TQ|PH|SD\rangle = 0, \\ \langle SD|PH|TQ\rangle = \langle SD|H|TQ\rangle. \end{cases} \tag{3.19}$$

$$\text{SCC}: \begin{cases} \langle TQ|\frac{1}{2}(PH + HP)|SD\rangle = \frac{1}{2}\langle TQ|H|SD\rangle, \\ \langle SD|\frac{1}{2}(PH + HP)|TQ\rangle = \frac{1}{2}\langle SD|H|TQ\rangle. \end{cases} \tag{3.20}$$

The effective structure of $\mathbf{H}$ is illustrated for both theories in Figure 3.2. In comment to a paper by Pedersen and Koch,[57] Moszynski and coworkers pointed out[93] that the wrong factor $1/2$ appears both below and above the diagonal. We consider this a compromise between the correct and completely ignored elements, respectively, over and under the diagonal in the traditional approach.

FIGURE 3.2: *Matrix elements of H in the traditional and symmetric theories. The relative sizes as compared to the exact* $\langle\varphi|H|\psi\rangle$ *are listed.*

As anticipated, the symmetric theory is Hermitian. While the energy is clearly real, $E_{\mathrm{scc}} = \Re e(\overline{H}_{00} + \boldsymbol{\xi}^\dagger \boldsymbol{\tau})$, the structure of the expectation value is of greater importance:

$$\langle H \rangle_{\mathrm{scc}} = \frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | P | \psi \rangle}, \quad \mathcal{H} = \frac{1}{2}(PH + HP). \tag{3.21}$$

One may accurately conceive of $\mathcal{H}$ as the symmetric coupled cluster approximation of H. In this regard, since $P|_{\mathscr{F}_\Lambda} = \mathbb{I}$, it may similarly be considered an approximation of $\mathbb{I}$. The fact that $\mathcal{H} = \mathcal{H}^\dagger$ and $P = P^\dagger$ is what ensures the reality of the excitation energies and the consistency of the linear response function (see Chapter 4). In contrast, traditional coupled cluster theory is not Hermitian, as was discussed in detail in Section 2.3.

At the singles and doubles truncation of T, both theories turn out to exhibit a computational scaling of N^6, where N is the number of electrons. Since an explanation of this fact requires the discussion of vectors, matrices and transformations introduced later (see Section 3.2.3), the reader may wish to continue on and return to this paragraph at a later point. Let us denote by v and o the number of virtual and occupied orbitals, respectively. By an order analysis we find that the initialization of $\mathbf{q}$ is done in $O(v^2 o^2)$ operations and that the transformations by $\mathbf{Q}, \mathbf{Q}^\dagger$ and $\mathbf{S}$ scale no worse than $O(v^3 o)$. The remaining terms in the symmetric amplitude equations—the initialization of $\boldsymbol{\eta}$ and $\boldsymbol{\xi}$ as well as the transformation by $\mathbf{A}$—exhibits no more than a sixth power combined occupied-virtual order (see Koch *et al.*[94]). Evaluating the right-hand-side of Equation (3.36) therefore requires $O(N^6)$ number of operations.

A major difference of the two approaches is that the symmetric formulation is *not* size-extensive: the energy of two non-interacting fragments A and

B is not necessarily equal to $E_A + E_B$. We postpone a proper explanation of this fact to Section 3.3.7, where the issue will be discussed further in light of the obtained results. A summary of the properties discussed in this section is provided in Table 3.1.

TABLE 3.1: *Theoretical properties of the symmetric and traditional theories. The two last entries relate to the singles and doubles truncation of T.*

	Symmetric	Traditional
Hermitian	Yes	No
Size-extensive	No	Yes
Exact in the limit	Yes	Yes
Computational scaling	N^6	N^6
Triples and quadruples	1/2, 1/2	1, 0

3.2. IMPLEMENTATION

We present a closed-shell implementation of symmetric coupled cluster theory at the singles and doubles level of theory (SCCSD) written in the Dalton[95] quantum chemistry program. Dalton's DIIS (see Pulay[96,97]) CCSD implementation was adapted to solve the symmetric amplitude equations.

3.2.1. Definitions in closed-shell coupled cluster theory

We have applied the standard definitions for the dual space and the cluster operator (see p. 685–698 in Helgaker *et al.*[1]), restricting the excitation operators to be operators of singlet symmetry. That is,

$$\mathscr{F}_\Lambda = \text{span} \{|\text{R}\rangle, E_{ai}|\text{R}\rangle, E_{ai}E_{bj}|\text{R}\rangle\}_{aibj}, \tag{3.22}$$

$$T = \sum_{ai} t_i^a E_{ai} + \frac{1}{2} \sum_{aibj} t_{ij}^{ab} E_{ai}E_{bj}, \tag{3.23}$$

where the $E_{ai} = a_{a\alpha}^\dagger a_{i\alpha} + a_{a\beta}^\dagger a_{i\beta}$ are termed *singlet excitation operators*.

3.2.2. The energy and amplitude equations revisited

In the present section we derive the energy and symmetric amplitude equations using bases that are standard in coupled cluster theory. Although the derivation follows the same steps as in Section 3.1.4, they become more involved due to presence of non-orthogonal bases for $\mathscr{F}_\Lambda$. A few preliminary remarks are therefore necessary.

We adopt the notation introduced in Equation (2.10), where we make the identifications $\tau_i^a = E_{ai}$ and $\tau_{ij}^{ab} = E_{ai}E_{bj}$. In terms of these vectors we may

define the *elementary basis* as

$$\mathcal{M} = \left\{ |\mathrm{R}\rangle, \left| \begin{matrix} a \\ i \end{matrix} \right\rangle, \left| \begin{matrix} ab \\ ij \end{matrix} \right\rangle \right\}_{aibj}. \tag{3.24}$$

It is useful to use this basis in combination with a basis that is bi-orthogonal or bi-orthonormal with respect to $\mathcal{M}$. The *bi-orthogonal* $\overline{\mathcal{M}}$ and *bi-orthonormal bases* $\widetilde{\mathcal{M}}$ serve this purpose:

$$\overline{\mathcal{M}} = \left\{ |\mathrm{R}\rangle, \left| \begin{matrix} \overline{a} \\ i \end{matrix} \right\rangle, \left| \begin{matrix} \overline{ab} \\ ij \end{matrix} \right\rangle \right\}_{aibj}, \tag{3.25}$$

$$\widetilde{\mathcal{M}} = \left\{ |\mathrm{R}\rangle, \left| \begin{matrix} \widetilde{a} \\ i \end{matrix} \right\rangle, \left| \begin{matrix} \widetilde{ab} \\ ij \end{matrix} \right\rangle \right\}_{aibj}, \tag{3.26}$$

where

$$\left| \begin{matrix} \overline{a} \\ i \end{matrix} \right\rangle = \frac{1}{2} \left| \begin{matrix} a \\ i \end{matrix} \right\rangle, \quad \left| \begin{matrix} \overline{ab} \\ ij \end{matrix} \right\rangle = \frac{1}{3} \left| \begin{matrix} ab \\ ij \end{matrix} \right\rangle + \frac{1}{6} \left| \begin{matrix} ab \\ ji \end{matrix} \right\rangle, \tag{3.27}$$

$$\left| \begin{matrix} \widetilde{a} \\ i \end{matrix} \right\rangle = \left| \begin{matrix} \overline{a} \\ i \end{matrix} \right\rangle, \quad \left| \begin{matrix} \widetilde{ab} \\ ij \end{matrix} \right\rangle = \frac{1}{1 + \delta_{ai,bj}} \left| \begin{matrix} \overline{ab} \\ ij \end{matrix} \right\rangle. \tag{3.28}$$

Observe that $\mathcal{M}$, $\overline{\mathcal{M}}$, and $\widetilde{\mathcal{M}}$ are all bases of the dual space $\mathscr{F}_\Lambda$. Denoting the respective determinants by $|\mu\rangle$, $|\overline{\mu}\rangle$, and $|\widetilde{\mu}\rangle$, the projection operator P onto $\mathscr{F}_\Lambda$ assumes the following three forms:

$$P = \sum_{\mu \geq 0} |\mu\rangle\langle\widetilde{\mu}| = \sum_{\mu\nu \geq 0} |\mu\rangle \widetilde{S}_{\mu\nu} \langle\nu| = \sum_{\mu\nu \geq 0} |\widetilde{\mu}\rangle S_{\mu\nu} \langle\widetilde{\nu}|. \tag{3.29}$$

We have introduced the overlap matrices $\mathbf{S}$ and $\widetilde{\mathbf{S}}$ associated with the bases $\mathcal{M}$ and $\widetilde{\mathcal{M}}$. They are matrix inverses of each other (*i.e.* $\mathbf{S}\,\widetilde{\mathbf{S}} = \mathbf{I}$).

We express quantities for which elements in $\widetilde{\mathcal{M}}$ enters as the bra state and $\mathcal{M}$ as the ket state with an added '$\sim$', for instance $\widetilde{Q}_{\mu\nu} = \langle\widetilde{\mu}|e^T|\nu\rangle$. All matrices and vectors will be denoted with the same letters as before. By introducing the appropriate form of P we can write

$$\langle\psi|P|\psi\rangle = 1 + \widetilde{\mathbf{q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}}, \tag{3.30}$$

$$\langle\psi|HP|\psi\rangle = (1 + \widetilde{\mathbf{q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}}) \overline{H}_{00}^\dagger + \widetilde{\boldsymbol{\xi}}^\dagger \widetilde{\mathbf{Q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}}. \tag{3.31}$$

Upon letting $\overline{\boldsymbol{\tau}} = \widetilde{\mathbf{Q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}}/(1 + \widetilde{\mathbf{q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}})$ and $N = 1 + \widetilde{\mathbf{q}}^\dagger \mathbf{S}\,\widetilde{\mathbf{q}}$ we find the energy

$$\langle H \rangle_{\mathrm{scc}} = \Re e\,(\overline{H}_{00} + \widetilde{\boldsymbol{\xi}}^\dagger \overline{\boldsymbol{\tau}}). \tag{3.32}$$

Finally, by partial differentiation of the relevant terms,

$$\frac{\partial}{\partial t_\mu^*}\langle\psi|P|\psi\rangle = (\widetilde{\mathbf{Q}}^\dagger\mathbf{S}\,\widetilde{\mathbf{q}})_\mu, \tag{3.33}$$

$$\frac{\partial}{\partial t_\mu^*}\langle\psi|HP|\psi\rangle = (N\,\boldsymbol{\eta} + (\widetilde{\mathbf{A}}^\dagger + E_{\text{scc}})\widetilde{\mathbf{Q}}^\dagger\mathbf{S}\,\widetilde{\mathbf{q}})_\mu, \tag{3.34}$$

$$\frac{\partial}{\partial t_\mu^*}\langle\psi|PH|\psi\rangle = (\widetilde{\mathbf{Q}}^\dagger\mathbf{S}\,\widetilde{\mathbf{q}}\,\overline{H}_{00} + \widetilde{\mathbf{Q}}^\dagger\mathbf{S}\widetilde{\mathbf{Q}}\,\widetilde{\boldsymbol{\xi}})_\mu, \tag{3.35}$$

we obtain the symmetric singles and doubles amplitude equations:

$$\mathbf{0} = \boldsymbol{\eta} + \widetilde{\mathbf{A}}^\dagger\overline{\boldsymbol{\tau}} + \alpha\,\overline{\boldsymbol{\tau}} + \frac{1}{N}\,\widetilde{\mathbf{Q}}^\dagger\mathbf{S}\,\widetilde{\mathbf{Q}}\,\widetilde{\boldsymbol{\xi}}. \tag{3.36}$$

3.2.3. Matrices, vectors, and transformations

We restrict ourselves to the listing of objects which are unique to SCCSD and refer the reader elsewhere for other quantities. The reader may consult Koch *et al.*[94] for $\boldsymbol{\eta}$ and $\widetilde{\mathbf{A}}$ and Scuseria *et al.*[98] for $\widetilde{\boldsymbol{\xi}}$. Moreover, $\overline{H}_{00}$ and N are given in Helgaker *et al.*,[1] p. 687 and 693. The remaining non-vanishing matrix elements are

$$S_{ai,ck} = 2\,\delta_{ai,ck}, \tag{3.37}$$

$$S_{aibj,ckdl} = 2\,\mathscr{P}_{ij}^{ab}(2\,\delta_{aibj,ckdl} - \delta_{ajbi,ckdl}), \tag{3.38}$$

$$\widetilde{q}_{ai} = t_i^a, \tag{3.39}$$

$$\widetilde{q}_{aibj} = \Delta_{ai,bj}^{-1}\,\mathscr{T}_{ij}^{ab}, \tag{3.40}$$

$$\widetilde{Q}_{ai,ck} = \delta_{ai,ck}, \tag{3.41}$$

$$\widetilde{Q}_{aibj,ckdl} = \Delta_{ai,bj}^{-1}\,\mathscr{P}_{ij}^{ab}\delta_{aibj,ckdl}, \tag{3.42}$$

$$\widetilde{Q}_{aibj,ck} = \Delta_{ai,bj}^{-1}\,(t_j^b\,\delta_{ai,ck} + t_i^a\,\delta_{bj,ck}), \tag{3.43}$$

where $\Delta_{ai,bj} = 1 + \delta_{ai,bj}$, $\mathscr{T}_{ij}^{ab} = t_{ij}^{ab} + t_i^a t_j^b$, and $\mathscr{P}_{ij}^{ab}$ permutes the ordered index pair (a,i) with (b,j).

The transformations in the symmetric amplitude equations are performed using restricted matrix product indices. This is due to the assumed form

$$P = |\text{R}\rangle\langle\text{R}| + \sum_{ai}\left|{a\atop i}\right\rangle\left\langle{\widetilde{a}\atop i}\right| + \sum_{ai\geq bj}\left|{ab\atop ij}\right\rangle\left\langle{\widetilde{ab}\atop ij}\right|, \tag{3.44}$$

which implies that matrix transformations $\mathbf{y} = \mathcal{O}\,\mathbf{x}$ run over restricted doubles indices. It is convenient to remove this restriction by use of $|{ab\atop ij}\rangle = |{ba\atop ji}\rangle$:

$$y_\mu = \sum_{ai}\mathcal{O}_{\mu,ai}\,x_{ai} + \frac{1}{2}\sum_{aibj}\mathcal{O}_{\mu,aibj}\,x_{aibj}\,\Delta_{ai,bj}. \tag{3.45}$$

For $\mathbf{y} = \mathbf{S}\,\mathbf{x}$ we find

$$y_{ai} = 2\,x_{ai}, \tag{3.46}$$

$$y_{aibj} = 2\,\Delta_{ai,bj}(2\,x_{aibj} - x_{ajbi}). \tag{3.47}$$

For $\mathbf{y} = \widetilde{\mathbf{Q}}\,\mathbf{x}$,

$$y_{ai} = x_{ai}, \tag{3.48}$$

$$y_{aibj} = \Delta^{-1}_{ai,bj}(t^b_j\,x_{ai} + t^a_i\,x_{bj}) + x_{aibj}. \tag{3.49}$$

And for $\mathbf{y} = \widetilde{\mathbf{Q}}^\dagger\,\mathbf{x}$,

$$y_{ai} = x_{ai} + \sum_{ck} t^{c\,*}_k\,x_{aick}, \tag{3.50}$$

$$y_{aibj} = x_{aibj}. \tag{3.51}$$

3.3. Results and discussion

We present a computational study on a series of small molecular systems. Energy calculations are performed in a restricted region of the potential energy surface in clusters of ammonia molecules, $(\mathrm{NH_3})_n$, with $n = 1, 2, 3, 4, 5$, and in $\mathrm{H_2O}$, where the symmetric stretching motion is studied. Full potential energy curves are presented for the diatomic systems $\mathrm{Be_2}$, $\mathrm{N_2}$, CO, HF, and LiH. A size-extensivity study is moreover performed, where the energies of a set of widely separated dimers are compared with their monomer values. Throughout, particular attention is directed toward analyzing the similarities and differences of the symmetric and traditional theories.

We have made use of Dunning's[99] correlation consistent cc-pVDZ basis set in all calculations. As points of reference, CCSD and CCSD(T) energies are computed in addition to SCCSD, and comparisons with exact full configuration interaction (FCI) energies are made whenever available. For the $\mathrm{Be_2}$ molecule we present our own FCI calculations, while the $\mathrm{N_2}$ values are taken from Chan *et al.*[59] and those for $\mathrm{H_2O}$ from Olsen *et al.*[100] Moreover, CISD calculations were conducted to quantify size-extensivity errors in the non-interacting dimers and the series of ammonia clusters. We employed the 2016 release of Dalton[95] to perform coupled cluster and full configuration interaction calculations, and the Q-Chem 4 program package[101] for truncated configuration interaction calculations.

3.3.1. Interaction energies in ammonia clusters

The geometries of the studied ammonia clusters are shown in Figure 3.3. Each individual $\mathrm{NH_3}$ monomer's N−H bond length and H−N−H angle is set to 1.02 Å and 109.47°, respectively. We define the coordinate R_* as the

distance separating the nitrogen atoms from an arbitrarily fixed central nitrogen, labeled by a '∗' in Figure 3.3. The internal geometry of each monomer is kept fixed with respect to R_*. Precise coordinates are given for $R_* = 3$ Å in Appendix B. Other R_* inputs may be obtained from these coordinates by *e.g.* the Avogadro editor.[102]

(A) *(NH₃)₂.*

(B) *(NH₃)₃.*

(C) *(NH₃)₄.*

(D) *(NH₃)₅.*
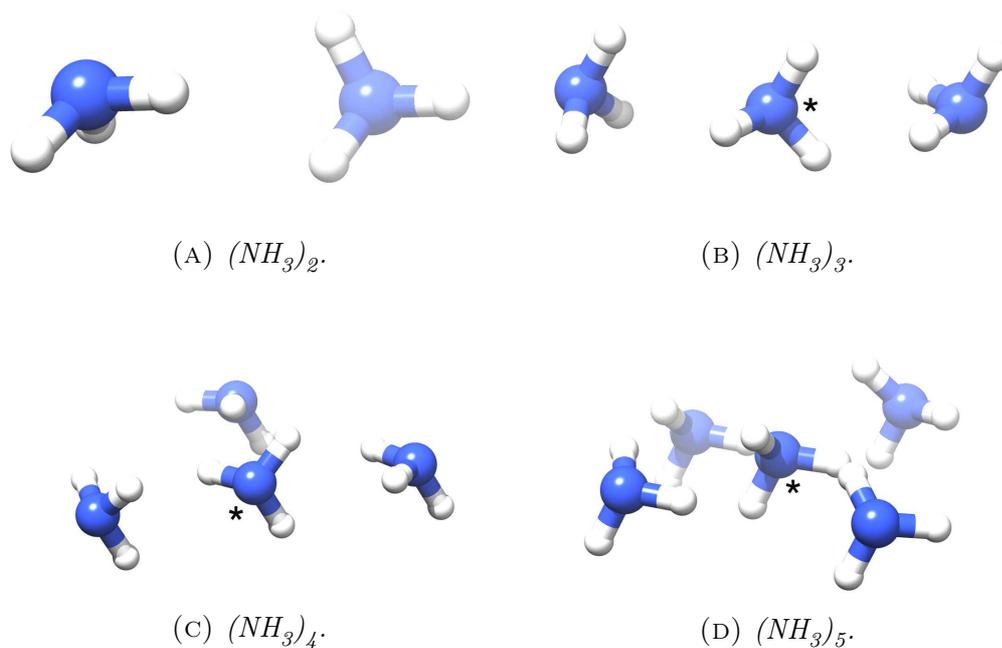
FIGURE 3.3: *The geometries of the ammonia clusters $(NH_3)_n$, $n = 2, 3, 4, 5$. All the nitrogen atoms are separated from the nitrogen labeled with a '∗' by three ångströms. The illustrations were produced with UCSF Chimera.*[103]

We define the *interaction energy* as the energy $E(R_*)$ at some distance R_* minus the energy at the dissociation limit $E(\infty)$. An adequate approximation of the dissociation limit is obtained at a distance of 40 ångströms:

$$\text{Interaction energy}(R_*) = E(R_*) - E(40 \text{ Å}). \tag{3.52}$$

For comparative purposes we inspect the *equilibrium distance* R_{eq} and the *well depth* $D_0 = E(40 \text{ Å}) - E(R_{\text{eq}})$, where it is understood that R_{eq} is the distance at which $E(R_*)$ reaches a minimum. The equilibrium distances were obtained by polynomial interpolation of the three points of lowest energy (see *e.g.* p. 805–809 in Kreyszig[104]).

The interaction energy curves are shown in Figure 3.4. Associated equilibrium distances and well depths are given in Table 3.2. Observe that SCCSD is in excellent agreement with CCSD for the dimer and trimer. In $(NH_3)_4$, SCCSD is in practice equivalent to CCSD(T). Large differences are however

obtained for the $(NH_3)_5$ cluster, where SCCSD provides over- and under-estimates of D_0 and R_{eq}, respectively. Reading off Table 3.2, note in particular the increasing difference between SCCSD and CCSD with respect the number of NH_3 fragments.

TABLE 3.2: *Equilibrium distances and well depths of the $(NH_3)_n$ clusters. The energy minima were estimated by interpolation of the three points of lowest energy.*

	R_{eq} [Å]				D_0 [mE$_h$]			
	CISD	CCSD	SCCSD	CCSD(T)	CISD	CCSD	SCCSD	CCSD(T)
$(NH_3)_2$	3.61	3.57	3.57	3.53	1.15	1.37	1.39	1.58
$(NH_3)_3$	3.52	3.48	3.47	3.44	3.92	4.68	4.84	5.17
$(NH_3)_4$	3.64	3.58	3.56	3.54	4.67	6.12	6.82	6.90
$(NH_3)_5$	3.63	3.55	3.46	3.52	6.33	8.59	12.90	9.67

3.3.2. The symmetric stretching motion of the water molecule

In all calculations, we fixed the H−O−H angle to 110.6°. The O−H distance is varied in multiples of the equilibrium distance $R_e = 1.84345$ a.u. Precise cartesian coordinates are given in Olsen *et al.*[100] See Table 3.3 for the obtained CCSD, SCCSD, and FCI energies at O−H bond lengths R_e, $1.5\,R_e$, $2.0\,R_e$, and $2.5\,R_e$. Close to the equilibrium length (at R_e and $1.5\,R_e$), SCCSD, CCSD, and FCI are in close agreement (< 0.01 hartrees). In contrast, considerable differences are found at $2.0\,R_e$, where the SCCSD energy is slightly below the FCI energy and CCSD lies well above it (by 0.02 hartrees). We were not able to converge the symmetric amplitude equations at O−H bond lengths greater than $2.0\,R_e$.

TABLE 3.3: *Energies along the symmetric stretching motion of H_2O. The length of the O−H bonds are multiples of $R_e = 1.84345$ a.u. The FCI numbers are from Olsen* et al.[100] *Energies are listed in hartrees. A '−' means that we were not able to converge the equations.*

	R_e	$1.5\,R_e$	$2.0\,R_e$	$2.5\,R_e$
CCSD	-76.23812	-76.06230	-75.92963	-75.89768
SCCSD	-76.23823	-76.06303	-75.95546	−
FCI	-76.24186	-76.07235	-75.95167	-75.91800
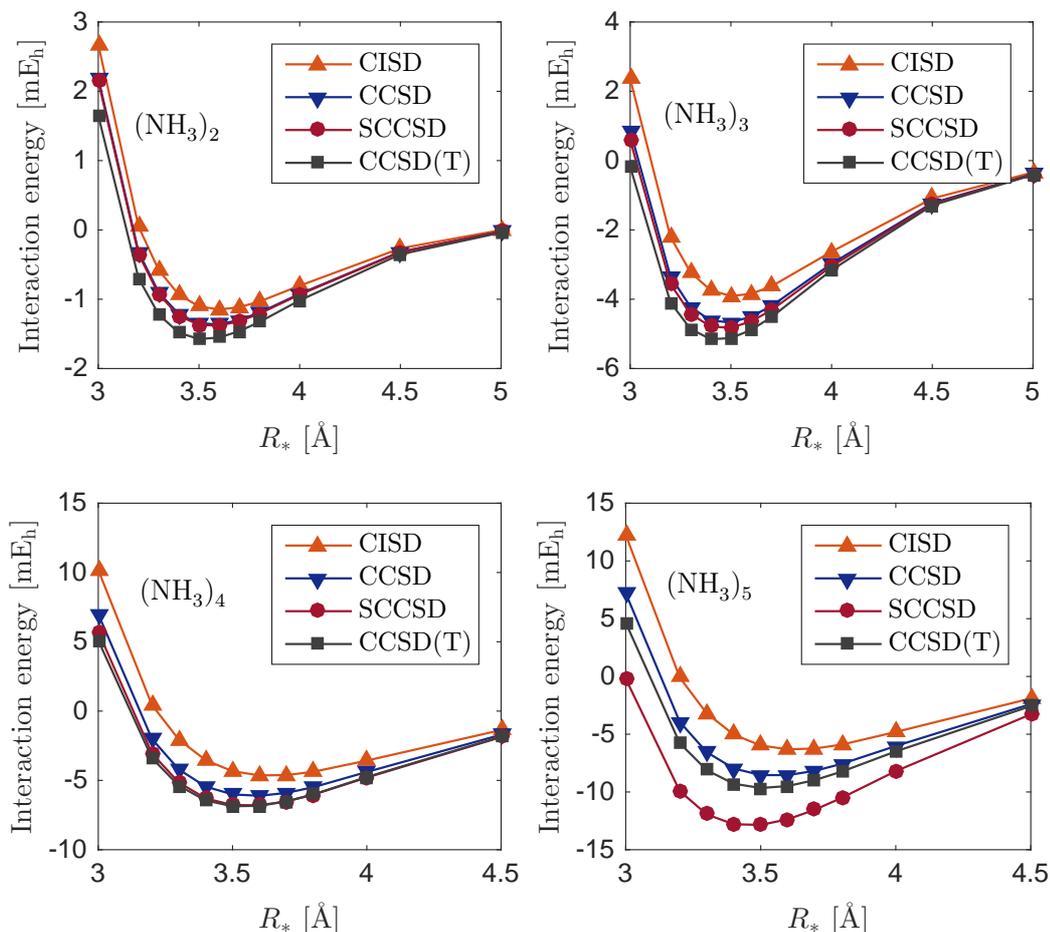
FIGURE 3.4: *Interaction energies of the ammonia clusters $(NH_3)_n$. Each curve is labeled with the identity of the cluster. The interaction energies are given in millihartrees (mE_h).*

3.3.3. Diatomic potential energy curves

We present potential energy curves of Be_2, N_2, CO, HF, and LiH, see Figures 3.5 and 3.6. Predicted equilibrium distances are given in Table 3.4. Observe that the coupled cluster curves incorrectly predicts a repulsive potential for Be_2 (see Figure 3.6). The failure to describe Be_2 is well known for single-reference approaches and has been attributed to the multireference nature of the bond.[105] We find that SCCSD and CCSD predict identical equilibrium bond lengths (within 0.1 picometers, see Table 3.4). Their energy curves are moreover seen to be indistinguishable close to the equilibrium geometry (see Figure 3.5). We were not able to converge the symmetric amplitude equations for $R > 3.2$ a.u. in N_2 nor for $R > 3.0$ a.u. in the CO molecule.

TABLE 3.4: *Equilibrium bond lengths of N_2, CO, HF, and LiH. Experimental values are from Helgaker* et al.[1] *All distances are in picometers. We calculated $E(R)$ for $R = 0.01 \cdot k$ atomic units for integers k (which translates to a separation of 0.529 picometers) and performed interpolation on the three points of lowest energy.*

	CCSD	SCCSD	CCSD(T)	Experiment
N_2	111.2	111.3	111.8	109.8
CO	113.8	113.9	114.4	112.8
HF	91.9	91.9	92.0	91.7
LiH	161.5	161.5	161.5	–

3.3.4. Size-extensivity study on widely separated dimers

A model is not size-extensive if the energy of a widely separated dimer (separated by a sufficient distance to be non-interacting) does not equal the sum of the energies of the individual fragments. We inspect

$$E_{\text{error}} = E_{\text{dimer}} - 2\,E_{\text{monomer}}, \tag{3.53}$$

comparing its size in symmetric coupled cluster and configuration interaction theories. The correlation energy of the latter model scales as $\sqrt{N}$, where N is the number of electrons.[8] Its energy therefore degenerates to the reference energy for large N.[1]

TABLE 3.5: *Size-extensivity errors in widely separated dimers. The error is calculated as $E_{\text{error}} = E_{\text{dimer}} - 2\,E_{\text{monomer}}$. All energies are given in hartrees.*

	E_{monomer}		E_{dimer}		E_{error}	
	SCCSD	CISD	SCCSD	CISD	SCCSD	CISD
Be	-14.6174	-14.6174	-29.2350	-29.2277	-0.0002	0.0071
H_2O	-76.2382	-76.2298	-152.4775	-152.4410	-0.0010	0.0186
NH_3	-56.3996	-56.3908	-112.8005	-112.7612	-0.0013	0.0203
CH_2O	-114.1555	-114.1342	-228.3164	-228.2285	-0.0054	0.0400

See Table 3.5 for the obtained errors in the beryllium, water, ammonia, and formaldehyde dimers. The monomers are separated by 100 ångströms in all calculations. Water and ammonia geometries are as given in the preceding sections. The O−C−H and H−C−H angles in formaldehyde (H_2CO) are set to 120° and the C−O and C−H bond lengths to 1.07 and 1.032 ångströms, respectively. Reading off Table 3.5 we see that the error of SCCSD is smaller than CISD in all the studied systems. Whereas the error is *thirty-five* times smaller for the the beryllium dimer, this factor is reduced to *twenty* for the
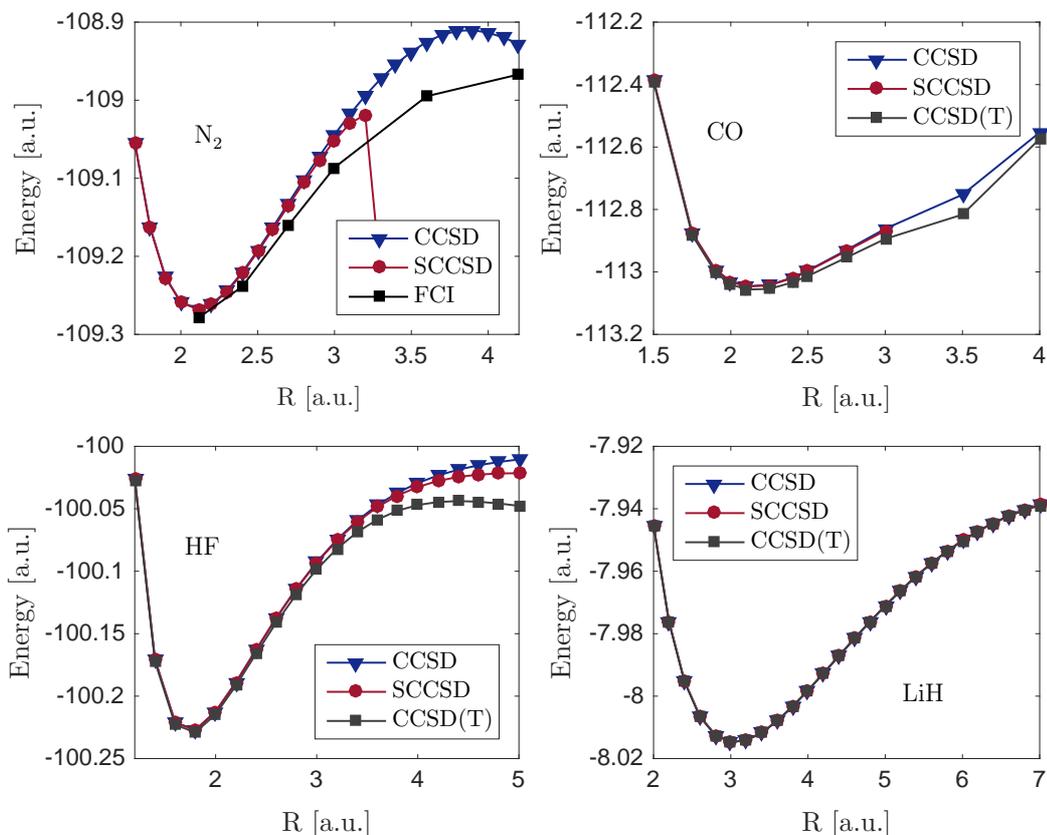
FIGURE 3.5: *Potential energy curves of N₂, CO, HF, and LiH.*

water dimer and to *eight* in the formaldehyde dimer. Extrapolating the data, we expect the error of SCCSD to exceed that of CISD for large systems. As the sign of the error differs in two methods, the SCCSD energy will grow rather than reduce to the energy of the reference $|R\rangle$.

3.3.5. Contributions to the energy in the clusters and diatomics

We investigate in this section the contributions to the symmetric coupled cluster energy. Let us write $E_{scc} = E_0 + \delta$, where $E_0 = \langle R|e^{-T}He^T|R\rangle$ and the *energy shift* $\delta = \overline{\tau}^\dagger \xi$ are assumed to be real numbers. In LiH, Be₂, and $(NH_3)_n$, both δ and the norms of $\overline{\tau}$ and ξ remain approximately constant with respect to the distance coordinates R and R_*. See Table 3.7 for their contributions at selected geometries. Variations of δ with the bond distance R are observed in N₂, CO, HF, and LiH. This variation is shown in Figure 3.7. A notable feature is that δ increases with the length of the bond (except in LiH). In the three-electron LiH system, δ is of the order 10^{-5} hartrees and therefore exerts no appreciable influence on the energy. For HF we list in Table 3.6 the contributions to the energy as well as the norm of the
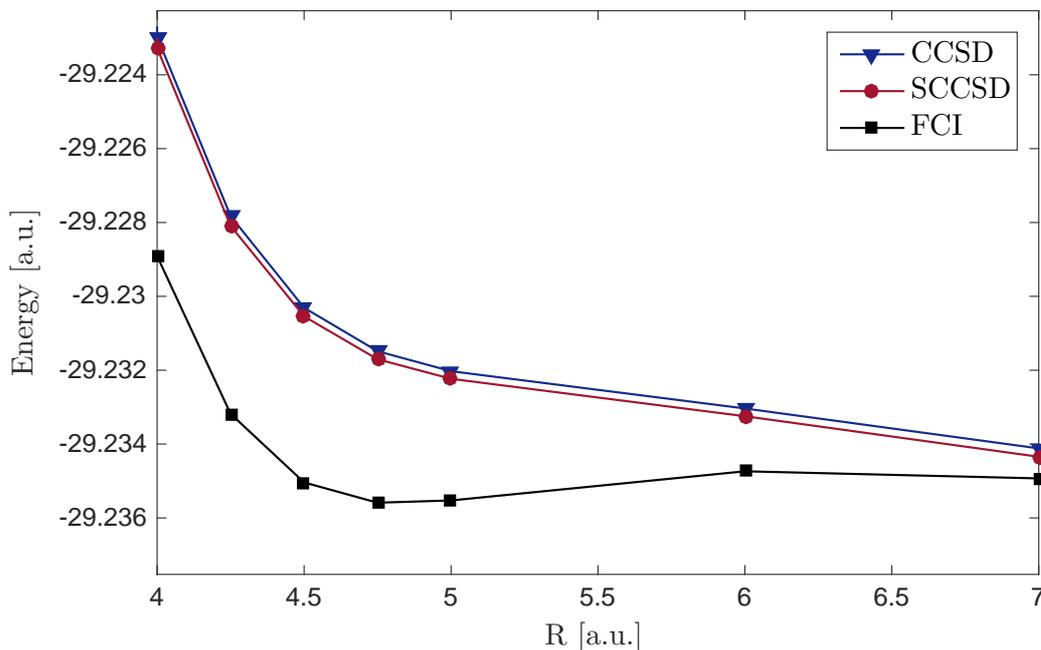
FIGURE 3.6: *The potential energy curve of Be$_2$.*

amplitudes **t**. Similar numbers are obtained for the N$_2$ and CO systems. We omit them for brevity.

This completes our survey of the results of the computational study. In the upcoming two sections, we turn to discuss in detail the performance and limitations of symmetric coupled cluster theory.

TABLE 3.6: *Contributions to the symmetric coupled cluster energy in HF. Listed is also the norm of the amplitudes **t**. The bond lengths R are in atomic units. The norms of $\overline{\tau}$ and **t** are dimensionless. All remaining values are listed in hartrees.*

R	E_0	$\|\boldsymbol{\xi}\|$	$\|\overline{\boldsymbol{\tau}}\|$	δ	E_{sccsd}	$\|\mathbf{t}\|$
2	-100.2166	0.0166	0.4219	0.0042	-100.2124	0.1661
3	-100.1017	0.0390	0.6528	0.0085	-100.0931	0.3083
4	-100.0562	0.1253	1.1404	0.0237	-100.0325	0.6067
5	-100.0704	0.2747	1.6407	0.0489	-100.0215	0.9286

3.3.6. The performance on small systems and the flexibility of the cluster operator

For molecular systems comprised of two electrons, the singles and doubles models are exact, *i.e.*, they equal each other and are indistinguishable from
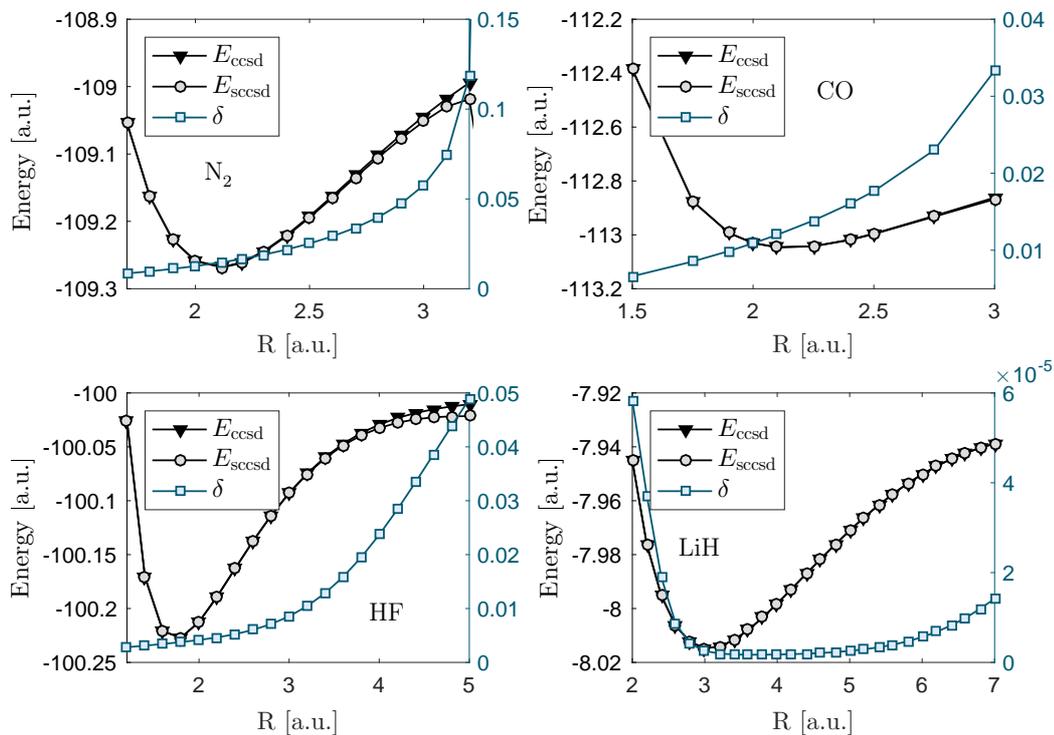
FIGURE 3.7: *The energy shift δ in N_2, CO, HF, and LiH. Axes on the right provide the values of δ while E_{ccsd} and E_{sccsd} may be read from the left axes.*

the full configuration interaction energies. There are however no guarantees that the theories should make the same predictions on systems with more than two electrons. For very small systems, *e.g.* the four-electron Be system, we obtain identical results both for the coupled cluster (SCCSD, CCSD) and configuration interaction (CISD) methods, see Tables 3.5 and 3.7.

Much more interesting are our findings for medium-sized molecules containing electrons numbering in the tens, twenties, and thirties. The equilibrium geometry energies of H_2O and NH_3, both ten-electron systems, show a CISD/CCSD difference of 80-90 millihartrees. The SCCSD/CCSD difference remains at a single millihartree (see Tables 3.3 and 3.7). Absolute energies can be misleading measures of quality, however. In nuclear dynamics, it is the shape of the potential energy surfaces that matter, and the shapes are determined by relative energies. Reassuring is therefore the close agreement of the symmetric and traditional theories in the Be_2, N_2, CO, and HF curves close to equilibrium distances (see Figure 3.5). Moreover, SCCSD and CCSD predicts nearly identical bond distances (see Table 3.4). We should mention that the poorer performance of CCSD(T) compared to the experimental values is attributable to error cancellation: a more complete basis set tends to shorten bond lengths, whereas a superior treatment of correlation lengthens them (see Tables 15.5–15.7 in Helgaker *et al.*[1]).

TABLE 3.7: *Contributions to the symmetric coupled cluster energy. We list values for $R_* = 3.5$ Å in the ammonia clusters $(NH_3)_n$, $R = 4.5$ a.u. in Be_2, and $R = 3.0$ a.u. in LiH. All values are given in hartrees, except the dimensionless norm of $\overline{\tau}$.*

	E_0	$\|\boldsymbol{\xi}\|$	$\|\overline{\boldsymbol{\tau}}\|$	δ	E_{sccsd}	E_{ccsd}
LiH	-8.0147	0.0004	0.2787	$<10^{-5}$	-8.0147	-8.0147
Be	-14.6174	0.0021	0.5908	$<10^{-5}$	-14.6174	-14.6174
Be$_2$	-29.2343	0.0135	0.7886	0.0037	-29.2305	-29.2303
NH$_3$	-56.4047	0.0167	0.4881	0.0051	-56.3996	-56.3995
(NH$_3$)$_2$	-112.8291	0.0588	0.7037	0.0272	-112.8019	-112.8003
(NH$_3$)$_3$	-169.2830	0.1315	0.9094	0.0731	-169.2099	-169.2031
(NH$_3$)$_4$	-225.7831	0.2497	1.1262	0.1585	-225.6246	-225.6039
(NH$_3$)$_5$	-282.4104	0.4960	1.4398	0.3448	-282.0657	-282.0059

We have shown that the symmetric and traditional theories predict identical energies *if* their amplitudes are identical (see Section 3.1.5). This would however imply that $\boldsymbol{\xi} = \mathbf{0}$, which is only true for the LiH and Be systems (see Table 3.7). The close agreement for larger systems is therefore unexpected. Consider *e.g.* the energy of the beryllium dimer:

$$E_{\text{cc}} = \langle \text{R} | e^{-T} H e^{T} | \text{R} \rangle = -29.2303 \text{ hartrees}$$
$$E_{\text{scc}} = \underbrace{\langle \text{R} | e^{-T} H e^{T} | \text{R} \rangle}_{-29.2343 \text{ hartrees}} + \underbrace{\overline{\boldsymbol{\tau}}^{\dagger} \boldsymbol{\xi}}_{0.0037 \text{ hartrees}} = -29.2305 \text{ hartrees} \qquad (3.54)$$

The traditional amplitudes make up T in the first expression; the symmetric amplitudes enter in the second. Similar numbers are found in N$_2$, CO, HF, and LiH. Moreover, although the theories start to differ in absolute energies for the ammonia dimer and trimer, the effect of this difference on the energy curves is limited (see Figure 3.4). These findings indicate that symmetric coupled cluster theory predicts different states (different t_μ) but identical energies (more precisely relative energies) in molecules comprised of about 10–30 electrons. The performance of the symmetric formulation is therefore *not* due to an equivalence with coupled cluster theory (*i.e.* $\boldsymbol{\xi} = 0$). It remains to be seen whether this interesting feature extends to its molecular properties and transition moments. Importantly, the cluster operator T exhibits a high degree of flexibility: the accuracy of the energies is not predicated on its amplitudes being determined by projection.

Large differences are observed for $R \gg R_{\text{eq}}$ in N$_2$ and HF (see Figure 3.5). Deviations are also seen in the stretching of H$_2$O (see Table 3.3). These discrepancies are best understood in relation to the often deteriorating quality of the restricted Hartree-Fock state $|\text{R}\rangle$ with bond length. It ceases to be

the leading term in the exact wavefunction, as demonstrated for H_2O and N_2 in Olsen *et al.*[1,100] and Chan *et al.*[59] Symptomatic of this issue is the growing norm of the amplitudes $\mathbf{t}$,[31] which acts to lessen the domination of $|\text{R}\rangle$ in the coupled cluster state $|\psi\rangle$ (see Table 3.6 for the HF values). The poor performance of the traditional and symmetric theories can therefore be attributed to the improper description provided by $|\text{R}\rangle$. Multireference theories are better suited in these regions (see Section 2.4.1). The differences between the symmetric and traditional theories are therefore significant only in regions where neither should be applied. They occur as $\boldsymbol{\xi}$ and $\overline{\boldsymbol{\tau}}$ grows with R, where the increase in $\overline{\boldsymbol{\tau}}$ is caused by $\mathbf{t}$ (see Table 3.6):

$$\overline{\tau}_\nu = \frac{\sum_\mu \langle \text{R} | e^{T^\dagger} | \mu \rangle \langle \mu | e^T | \nu \rangle}{1 + \sum_\mu \langle \text{R} | e^{T^\dagger} | \mu \rangle \langle \mu | e^T | \text{R} \rangle} = \frac{O(T)}{1 + O(T)}. \tag{3.55}$$

We emphasize that symmetric coupled cluster theory should not be applied in regions with significant multireference character, even if some have advocated alternative formulations in such regimes.[106] Its single-reference nature notwithstanding, we frequently encounter convergence issues and the occasional curve collapse in these regions. The reader is referred to Evangelista[107] for a critical assessment of alternative coupled cluster formulations in multireference regimes.

In $(NH_3)_n$, $n = 2, 3, 4$, SCCSD and CCSD give similar predictions and lie closer to CCSD(T) than CISD does (see Table 3.2 and Figure 3.4). For $(NH_3)_5$ the SCCSD predictions differ by a large margin from those of CCSD(T). This is clear indication that the symmetric and traditional theories will part ways for large systems, limiting the applicability of SCCSD to systems comprised of fewer than about 50 electrons. The non-extensivity of the method offers an explanation of the major changes observed from $(NH_3)_3$ to $(NH_3)_5$. Although scaling errors relate to infinitely separated and non-interacting fragments, they are thought to be detrimental for the model accuracy on large systems.[12,79] Moreover, the abrupt on-set of poor description may be understood from the size-extensivity study. Reading off Table 3.5 we see that, while the errors of SCCSD are smaller than those of CISD, the error increases at a higher rate. No abrupt change is observed in CISD for $(NH_3)_5$. Finally, the molecular size (in terms of the number of electrons) at which the SCCSD and CISD errors are expected to become comparable matches the on-set of poor description in the ammonia clusters.

We have found that $\boldsymbol{\xi}$ differs from zero in all but very small molecules. In systems containing fewer than thirty electrons, this difference mainly affects the amplitudes, not the energy. The changes in the block structure of $\mathbf{H}$ appear to be of little consequence (see Figure 3.2). We note that the shift δ allows T_1^3, T_1^4, T_2^2, and $T_1 T_2$ to contribute to the energy, however; projection implies that only T_1, T_1^2, and T_2 enters in the traditional coupled cluster energy (see Section 3.1.5). It is tempting to conjecture that the approximate

inclusion of higher-order excitations in SCCSD could offer improvements over CCSD. Observe that SCCSD is closer to CCSD(T) in $(NH_3)_3$ and $(NH_3)_4$ (see Table 3.2). It remains unclear whether these improvements are genuine or merely artefacts caused by size-extensivity errors. Further investigations will be necessary to settle the question. In any case, the additional dependence on disconnected contributions might explain why SCCSD is less robust than CCSD in multiference regimes (see Figure 3.5 and Section 2.4.1).

We must emphasize that many molecules of interest contain more than 50 electrons, and that symmetric coupled cluster theory therefore has limited applicability. Due to the accurate results obtained in small systems we remain optimistic that related formulations exist which exhibit more favorable scaling properties. It is thus of importance for future work that we understand why symmetric coupled cluster theory does not satisfy additive nor multiplicative separability. This is the purpose of the next section.

3.3.7. The loss of size-extensivity and the discrepancies of the symmetric and traditional theories on large systems

A general analysis of the symmetric amplitude equations becomes quite involved, and we will therefore treat in detail the special case of two separated non-interacting fragments which both satisfy the traditional amplitude equations if isolated (the beryllium dimer is a case in point, see Table 3.7). Below we analyze this particular case, showing the wavefunction and energy does not separate. After the analysis we give a brief consideration of the general case, before turning to identify the root cause of the loss of separability.

Recall that whenever the symmetric and traditional amplitudes are identical, the symmetric amplitude equations imply that $\overline{\boldsymbol{\tau}}$ must satisfy the multiplier equation (see Section 3.1.5):

$$\boldsymbol{\eta}^{\dagger} + \overline{\boldsymbol{\tau}}^{\dagger}\,\mathbf{A} = \mathbf{0}. \tag{3.56}$$

This statement can be understood by letting $\overline{t}_{\mu} = \overline{\tau}_{\mu} = (\mathbf{q}^{\dagger}\mathbf{Q})_{\mu}/(1+\mathbf{q}^{\dagger}\mathbf{q})$ in the expression for the traditional dual state $\langle\psi_{\Lambda}|$. Making use of the identity $P = |\mathrm{R}\rangle\langle\mathrm{R}| + \sum_{\mu}|\mu\rangle\langle\mu|$ and $P|_{\mathscr{F}_{\Lambda}} = \mathbb{I}$ we find

$$\langle\psi_{\Lambda}| = \langle\mathrm{R}| + \sum_{\mu}\overline{t}_{\mu}\langle\mu|e^{-T} = \frac{(\langle\mathrm{R}| + \sum_{\mu}q_{\mu}^{\dagger}\langle\mu|)}{1+\mathbf{q}^{\dagger}\mathbf{q}} = \frac{\langle\psi|P}{\langle\psi|P|\psi\rangle}, \tag{3.57}$$

where $|\psi\rangle$ is the coupled cluster state. The definitions of $\mathbf{q}$ and $\mathbf{Q}$ are given on p. 23. The symmetric and traditional theories thus differ unless the traditional dual $\langle\psi_{\Lambda}|$ is the projected state $P|\psi\rangle$ (ignoring normalization).

Now consider a compound system AB comprised of two non-interacting systems A and B, *e.g.* two widely separated beryllium monomers. We assume that the symmetric and traditional amplitudes are identical in the isolated A and B systems. By coupled cluster theory's size-extensivity we find that

$T = T_A + T_B$ satisfies $\boldsymbol{\xi} = \mathbf{0}$ by virtue of $\boldsymbol{\xi}_A = \boldsymbol{\xi}_B = \mathbf{0}$. The question to answer is whether the projection of

$$|\psi\rangle = e^T |\mathrm{R}\rangle = e^{T_A}|\mathrm{R}_A\rangle \otimes e^{T_B}|\mathrm{R}_B\rangle \tag{3.58}$$

is the traditional dual state $\langle\psi_\Lambda|$. This is not so if $\boldsymbol{\eta}^\dagger + \overline{\boldsymbol{\tau}}^\dagger\,\mathbf{A} \neq \mathbf{0}$, resulting in a change in the amplitudes (and hence the energies) of the compound system *vis-à-vis* the traditional theory.

Let us show that $\overline{\boldsymbol{\tau}}$ does not satisfy the multiplier equation. We denote by 'A', 'B', or 'AB' the blocks of $\mathbf{A}$, $\mathbf{Q}$, $\boldsymbol{\eta}$, and $\mathbf{q}$ which involve τ_μ in the respective subsystems, *e.g.* $\langle\mu_A\mathrm{R}_B|e^T|\nu_A\nu_B\rangle$ is an element of $\mathbf{Q}_{A,AB}$. It is not difficult to derive the block structures

$$\boldsymbol{\eta} = \begin{pmatrix} \boldsymbol{\eta}_A \\ \boldsymbol{\eta}_B \\ 0 \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} \mathbf{A}_{A,A} & 0 & \mathbf{A}_{A,AB} \\ 0 & \mathbf{A}_{B,B} & \mathbf{A}_{B,AB} \\ 0 & 0 & \mathbf{A}_{AB,AB} \end{pmatrix}, \tag{3.59}$$

and

$$\mathbf{q} = \begin{pmatrix} \mathbf{q}_A \\ \mathbf{q}_B \\ \mathbf{q}_{AB} \end{pmatrix}, \quad \mathbf{Q} = \begin{pmatrix} \mathbf{Q}_{A,A}^\dagger & 0 & \mathbf{Q}_{AB,A}^\dagger \\ 0 & \mathbf{Q}_{B,B}^\dagger & \mathbf{Q}_{AB,B}^\dagger \\ 0 & 0 & \mathbf{Q}_{AB,AB}^\dagger \end{pmatrix}^\dagger. \tag{3.60}$$

Note that $\boldsymbol{\eta}_A$, $\mathbf{q}_A$, $\mathbf{A}_{A,A}$, and $\mathbf{Q}_{A,A}$ equal the corresponding objects in the isolated A system (and likewise for B). By letting $\square$ denote arbitrary AB contributions, $\overline{\boldsymbol{\tau}}^\dagger = \mathbf{q}^\dagger\mathbf{Q}/(1 + \mathbf{q}^\dagger\mathbf{q})$ can be written

$$\overline{\boldsymbol{\tau}}^\dagger = \frac{1}{1 + \mathbf{q}_A^\dagger\mathbf{q}_A + \mathbf{q}_B^\dagger\mathbf{q}_B + \square}(\mathbf{q}_A^\dagger\mathbf{Q}_{A,A} + \square,\ \mathbf{q}_B^\dagger\mathbf{Q}_{B,B} + \square,\ \square). \tag{3.61}$$

It is evident that the A and B blocks of $\overline{\boldsymbol{\tau}}^\dagger\mathbf{A}$ are different from $\overline{\boldsymbol{\tau}}_A^\dagger\mathbf{A}_{A,A}$ and $\overline{\boldsymbol{\tau}}_B^\dagger\mathbf{A}_{B,B}$ in general. The identities $\boldsymbol{\eta}_i^\dagger + \overline{\boldsymbol{\tau}}_i^\dagger\mathbf{A}_{i,i} = \mathbf{0}$ are thus not sufficient to guarantee zeros in the A and B blocks of $\boldsymbol{\eta} + \overline{\boldsymbol{\tau}}^\dagger\mathbf{A}$. There are also a large number of non-zero contributions to the AB block of $\overline{\boldsymbol{\tau}}^\dagger\mathbf{A}$.

We conclude that $T \neq T_A + T_B$ and hence $|\psi\rangle \neq |\psi_A\rangle \otimes |\psi_B\rangle$. In turn this implies that $E \neq E_A + E_B$ because $E_{\mathrm{scc}} = \Re e\left(\overline{H}_{00} + \overline{\boldsymbol{\tau}}^\dagger\boldsymbol{\xi}\right)$ contains several AB terms for $\boldsymbol{\xi} \neq 0$ and $T = T_A + T_B + T_{AB}$. In the general analysis, the last row of $\mathbf{A}$ is non-zero, causing a further proliferation of AB terms that do not vanish in the symmetric amplitude equations.

A better understanding of the loss of extensivity is gained by considering the influence of the dual state in the symmetric and traditional approaches. Observe that the dependence on $\langle\psi_\Lambda|$ is only apparent in the traditional coupled cluster energy:

$$E_{\mathrm{cc}} = \langle\psi_\Lambda|H|\psi\rangle = \langle\mathrm{R}|\overline{H}|\mathrm{R}\rangle. \tag{3.62}$$

If the energy depended on $\langle\psi_\Lambda|$ the separability of T would not be sufficient to guarantee that $E_{\mathrm{cc}} = E_{\mathrm{cc}}^A + E_{\mathrm{cc}}^B$. The extensivity of the energy is predicated on the projection of the Schrödinger equation: the multiplier contributions in $\langle\psi_\Lambda| = \langle\mathrm{R}| + \sum_\mu \bar{t}_\mu\langle\mu|e^{-T}$ vanish for precisely this reason.

The dual state is more influential in the symmetric formulation. Let us write $E_{\mathrm{scc}} = \Re e\,\langle\psi_\Lambda|H|\psi\rangle$, where $\langle\psi_\Lambda| = \langle\mathrm{R}| + \sum_\mu \bar{\tau}_\mu^\dagger\langle\mu|e^{-T}$. Assuming that the energy is a real number we can write

$$E_{\mathrm{scc}} = \langle\mathrm{R}|\overline{H}|\mathrm{R}\rangle + \sum_\mu \bar{\tau}_\mu^\dagger \xi_\mu. \tag{3.63}$$

The dependence on $\overline{\boldsymbol{\tau}}$ and hence $\langle\psi_\Lambda|$ is evident. Its influence is restrained by $\boldsymbol{\xi}$, however, and thus non-existent for small systems (see Table 3.2). We have seen that the first contribution to E_{scc} is non-extensive by $T \neq T_A + T_B$ and the second by $\overline{\boldsymbol{\tau}}^\dagger \neq (\overline{\boldsymbol{\tau}}_A^\dagger\,\overline{\boldsymbol{\tau}}_B^\dagger\,\overline{\boldsymbol{\tau}}_{AB}^\dagger)$.

Although the independence of $\langle\psi_\Lambda|$ is not inherited by the traditional linear response functions and the associated transition moments, projection is still responsible for their correct scaling behavior. Consider for instance the right transition moment[19,64]

$$\begin{aligned}
\langle\psi_0|\mathcal{O}|\psi_k\rangle = {} &\sum_\mu\langle\psi_\Lambda|[\mathcal{O},\tau_\mu]|\psi\rangle c_{\mu k} \\
&+ \sum_{\mu\nu} c_\nu^{\mathcal{O}}(-\omega_k)\langle\psi_\Lambda|[[H,\tau_\nu],\tau_\mu]|\psi\rangle c_{\mu k},
\end{aligned} \tag{3.64}$$

where $\mathbf{A}\mathbf{c}_k = \omega_k\,\mathbf{c}_k$ and $(\mathbf{A} + \omega\,\mathbf{I})\mathbf{c}^{\mathcal{O}}(\omega) = \boldsymbol{\xi}^{\mathcal{O}}$ for $\xi_\mu^{\mathcal{O}} = \langle\mu|e^{-T}\mathcal{O}\,e^T|\mathrm{R}\rangle$. For two non-interacting subsystems A and B, the block structure of $\mathbf{A}$ implies

$$(-\mathbf{A} + \omega\,\mathbf{I})\begin{pmatrix}\mathbf{c}_A^{\mathcal{O}}(\omega)\\\mathbf{c}_B^{\mathcal{O}}(\omega)\\0\end{pmatrix} = \begin{pmatrix}\boldsymbol{\xi}_A^{\mathcal{O}}\\\boldsymbol{\xi}_B^{\mathcal{O}}\\0\end{pmatrix}, \quad \mathbf{A}\begin{pmatrix}\mathbf{c}_{kA}\\0\\0\end{pmatrix} = \omega_{kA}\begin{pmatrix}\mathbf{c}_{kA}\\0\\0\end{pmatrix}. \tag{3.65}$$

Substituting these compound system solutions ($\mathbf{c}^{\mathcal{O}}(\omega)$ and $\mathbf{c}_k$) in the transition element yields the desired size-intensivity: $\langle\psi_0|\mathcal{O}|\psi_k\rangle = \langle\psi_0^A|\mathcal{O}_A|\psi_k^A\rangle$. This relation is true because only the $\overline{\mathbf{t}}_A$ part of $\overline{\mathbf{t}} = (\overline{\mathbf{t}}_A\,\overline{\mathbf{t}}_B\,\overline{\mathbf{t}}_{AB})$ contributes to the transition moment. The extensivity of the linear response functions is valid for the same reason.[64]

The projection principle enters in the block structure of $\mathbf{A}$. It is worth noting that this structure is responsible for the relations of Equation (3.65). The most important blocks, $\mathbf{A}_{AB,A}$ and $\mathbf{A}_{AB,B}$, are zero because of projection:

$$\begin{aligned}
\langle\mu_A\mu_B|A|\nu_A\mathrm{R}_B\rangle &= A_{\mu_A\nu_A}^A\langle\mu_B|e^{-T_B}(H_B - E_B)e^{T_B}|\mathrm{R}_B\rangle = 0, \\
\langle\mu_A\mu_B|A|\mathrm{R}_A\nu_B\rangle &= \langle\mu_A|e^{-T_A}(H_A - E_A)e^{T_A}|\mathrm{R}_A\rangle A_{\mu_B\nu_B}^B = 0.
\end{aligned} \tag{3.66}$$

As correct scaling and projection are intimately connected, we deem it unlikely that a response theoretical extension of symmetric coupled cluster theory exhibits the proper scaling with system size.

Chapter 4

Symmetric equation of motion theory

4.1. Theory

4.1.1. The symmetric equation of motion *ansatz*

Recall that the amplitudes derived from the symmetric amplitude equations serve to define the state space $\mathscr{F}_\psi = \operatorname{span} \{e^T |\mathrm{R}\rangle, e^T |\mu\rangle\}_\mu$. In the traditional approach, the excited states are expanded in $\mathscr{F}_\psi$ (see Section 2.2.2). The symmetric equation of motion (SEOM) *ansatz* adds the restriction that the duals are their projections onto $\mathscr{F}_\Lambda$:

$$|\psi\rangle = e^T \left(\sum_{\mu \geq 0} c_\mu |\mu\rangle \right), \quad |\psi_\Lambda\rangle = P |\psi\rangle. \tag{4.1}$$

The cluster operator T is considered fixed and determined by the symmetric amplitude equations. It is understood that the sum runs over determinants $|\mu\rangle$ that belong to the dual space basis (*e.g.* $|\mathrm{R}\rangle$, $|_i^a\rangle$, and $|_{ij}^{ab}\rangle$ for $T = T_1 + T_2$).

The excited states are determined from the symmetric expectation value, as was done for the ground state in Chapter 3. The expansion coefficients c_μ are obtained by stationarity of $\langle H \rangle_{\mathrm{scc}}$ with respect to the c_μ^*:

$$\frac{\partial \langle H \rangle_{\mathrm{scc}}}{\partial c_\mu^*} = 0. \tag{4.2}$$

This is the *symmetric equation of motion equation*. It is solved for the c_μ, from which the excited states $|\psi_n\rangle$ are given by Equation (4.1). The discussion in Section 3.1.2 and Theorem A.1 imply that the Hellmann-Feynman theorem for $\langle H \rangle_{\mathrm{scc}}$ is valid for each excited state $|\psi_n\rangle$.

4.1.2. The full symmetric equation of motion equation

Due to presence of the reference coefficient in $|\psi\rangle$ we will frequently encounter matrix representations in bases that include $|\mathrm{R}\rangle$. To stay consistent with previously introduced notation, we reserve the notation $\mathcal{O}$ for reference-excluding representations, letting $\mathcal{O}^r$ denote matrices that include $|\mathrm{R}\rangle$. Recall that the symmetric coupled cluster approximation of H is

$$\mathcal{H} = \frac{1}{2}(PH + HP), \tag{4.3}$$

and that P is its approximation of the identity operator $\mathbb{I}$.

Deriving the symmetric equation of motion equation is most easily done by first deriving the required partial derivatives. As $\partial|\psi\rangle/\partial c_\mu = e^T |\mu\rangle$,

$$\frac{\partial}{\partial c_\mu^*}\langle\psi|P|\psi\rangle = \sum_{\nu\geq 0}\langle\mu|e^{T^\dagger}P\,e^T|\nu\rangle c_\nu = (\boldsymbol{P}^r\,\mathbf{c})_\mu, \qquad (4.4)$$

$$\frac{\partial}{\partial c_\mu^*}\langle\psi|\mathcal{H}|\psi\rangle = \sum_{\nu\geq 0}\langle\mu|e^{T^\dagger}\mathcal{H}\,e^T|\nu\rangle c_\nu = (\boldsymbol{\mathcal{H}}^r\,\mathbf{c})_\mu, \qquad (4.5)$$

where $\boldsymbol{P}^r$ and $\boldsymbol{\mathcal{H}}^r$ are matrix representations of P and $\mathcal{H}$ in the standard basis of $\mathscr{F}_\psi$. With these partial derivatives we immediately obtain

$$\frac{\partial\langle H\rangle_{\text{scc}}}{\partial c_\mu^*} = \frac{\partial}{\partial c_\mu^*}\frac{\langle\psi|\mathcal{H}|\psi\rangle}{\langle\psi|P|\psi\rangle} = \frac{(\boldsymbol{\mathcal{H}}^r\mathbf{c})_\mu - E\,(\boldsymbol{P}^r\,\mathbf{c})_\mu}{\langle\psi|P|\psi\rangle} = 0, \qquad (4.6)$$

where E is the energy of $|\psi\rangle$ as given by $\langle H\rangle_{\text{scc}}$. The solutions $\mathbf{c}_n$ are *eigenvectors* of the generalized Hermitian eigenvalue equation

$$\boldsymbol{\mathcal{H}}^r\,\mathbf{c}_n = E_n\,\boldsymbol{P}^r\,\mathbf{c}_n. \qquad (4.7)$$

Let us mention that this equation is valid for other choices of T, for instance the traditional cluster amplitudes. However, the ground state will not be the coupled cluster state $|\psi\rangle = e^T|\mathrm{R}\rangle$ for such T.

The symmetric coupled cluster state is a solution with the energy E from the ground state theory (see Chapter 3). Let us subtract $E\,\boldsymbol{P}^r\mathbf{c}_n$ from the equation and define the excitation energies as $\omega_n = E_n - E$. This results in the *full* symmetric equation of motion equation

$$\boldsymbol{\Omega}^r\,\mathbf{c}_n = \omega_n\,\boldsymbol{P}^r\,\mathbf{c}_n, \qquad (4.8)$$

where $\boldsymbol{\Omega}^r = \frac{1}{2}(\boldsymbol{P}^r\boldsymbol{A}^r + \boldsymbol{A}^{r\dagger}\boldsymbol{P}^r)$ is the representation of $\Omega = \mathcal{H} - E\,P$ in the standard basis of $\mathscr{F}_\psi$. We now turn to derive some of the theoretical properties of the method, delving into the equation's structure afterward.

4.1.3. The reality of the excitation energies and the completeness and approximate orthogonality of the eigenstates

A widely known and important property of Hermitian matrices is that they possess real eigenvalues and that there exists a basis composed of their eigenvectors.[71] Less well known are the properties of generalized Hermitian eigenvalue problems, matrix equations of the form

$$\mathbf{U}\,\mathbf{c}_n = \lambda_n\,\mathbf{V}\,\mathbf{c}_n, \qquad (4.9)$$

where $\mathbf{U} = \mathbf{U}^\dagger$ and $\mathbf{V} = \mathbf{V}^\dagger$ are Hermitian matrices. It turns out that if either $\mathbf{U}$ or $\mathbf{V}$ is positive-definite, there exists a basis of eigenvectors $\mathbf{c}_n$ such that

$$\mathbf{C}^\dagger\mathbf{U}\,\mathbf{C} = \mathrm{diag}(u_1,\ldots,u_m), \quad \mathbf{C}^\dagger\mathbf{V}\,\mathbf{C} = \mathrm{diag}(v_1,\ldots,v_m), \qquad (4.10)$$

where $\lambda_n = u_n/v_n$. By the Hermiticity of $\mathbf{C}^\dagger \mathbf{U}\, \mathbf{C}$ and $\mathbf{C}^\dagger \mathbf{V}\, \mathbf{C}$, $u_n, v_n \in \mathbb{R}$ and hence $\lambda_n \in \mathbb{R}$. A Hermitian generalized eigenvalue problem thus has real eigenvalues and a basis of eigenvectors if either of the matrices are positive definite (see Theorem 8.7.1 and Corollary 8.7.2 in Golub and Van Loan[108]). A review of available algorithms has been given by Bai and Demmel.[109]

Let us prove that the full symmetric equation of motion equation falls within this category by proving that $\boldsymbol{P}^r$ is positive definite. For an arbitrary vector $\mathbf{c}$ we have

$$\mathbf{c}^\dagger \boldsymbol{P}^r \mathbf{c} = \langle \psi, P\psi \rangle = \langle P\psi, P\psi \rangle \geq 0, \tag{4.11}$$

where we have let $|\psi\rangle = e^T \sum_{\mu \geq 0} c_\mu |\mu\rangle$ and used[67] $P = P^2$ and $P = P^\dagger$. This proves positive semi-definiteness. By the variational theorem,

$$\mathbf{c}^\dagger \boldsymbol{P}^r \mathbf{c} \geq \inf \sigma(\boldsymbol{P}^r), \tag{4.12}$$

and it suffices to show that zero is not an eigenvalue of $\boldsymbol{P}^r$. It is not difficult to see that $\boldsymbol{P}^r = \boldsymbol{Q}^{r\dagger} \boldsymbol{Q}^r$, where $Q^r_{\mu\nu} = \langle \mu | e^T | \nu \rangle$ is an invertible matrix with $(Q^r)^{-1}_{\mu\nu} = \langle \mu | e^{-T} | \nu \rangle$. As a product of invertible matrices, $\boldsymbol{P}^r$ must be invertible. But invertible matrices never have zero as an eigenvalue; it would contradict the eigenvalue condition $\det(\boldsymbol{P}^r - \lambda \mathbf{I}) = \det(\boldsymbol{P}^r) = 0$.[85] This finishes the proof.

Another attractive property of Hermitian matrices is that eigenvectors associated with different eigenvalues are orthogonal.[71] A weaker result appears in the generalized case. For $|\psi_n\rangle = e^T \sum_{\mu \geq 0} c_{n\mu} |\mu\rangle$ and $\mathbf{c}_n = (c_{n\mu})_\mu$,[109]

$$\mathbf{c}_m^\dagger\, \boldsymbol{P}^r\, \mathbf{c}_n = \langle P\psi_m, P\psi_n \rangle = 0. \tag{4.13}$$

To show this, let $\boldsymbol{P}^r$ act on $\mathbf{c}_n$ and $\mathbf{c}_m$ and use Equation (4.8). Whereas the excited states are not themselves orthogonal, their projections onto $\mathscr{F}_\Lambda$ are. The $\mathbf{c}_n$ are said to exhibit $\boldsymbol{P}^r$-orthogonality.

4.1.4. Transition moments and first-order molecular properties

The linear response functions are derived by substitution of the true states with those obtained from $\boldsymbol{\Omega}^r \mathbf{c}_n = \omega_n \boldsymbol{P}^r \mathbf{c}_n$. Our approach is best understood by re-writing the true response function as

$$\langle\langle A; V^\omega \rangle\rangle_\omega = -\langle \psi_0 | A\, R(\omega)\, V^\omega | \psi_0 \rangle - \langle \psi_0 | V^\omega\, R(-\omega)\, A | \psi_0 \rangle, \tag{4.14}$$

where the operator $R(\omega) = (H - E_0 - \omega)^{-1}$ is the *resolvent* of $H - E_0$.[67] We note that $\Omega = H - E_0$ in an untruncated expansion. The states $|\psi_n\rangle$ are here the eigenstates of H. To derive this expression of $\langle\langle A; V^\omega \rangle\rangle_\omega$, observe that

$$R(\omega) = \sum_n \frac{|\psi_n\rangle\langle\psi_n|}{\omega_n - \omega} \tag{4.15}$$

by the spectral theorem[67,110] $f(H) = \sum_n f(E_n) |\psi_n\rangle\langle\psi_n|$. We argue that the resolvent $R(\omega)$ has a unique analogue in our symmetric formalism, thereby determining the response function $\langle\langle A; V^\omega\rangle\rangle_\omega$.

In the following, we denote the states of symmetric equation of motion theory by $|\psi_n\rangle$. While $\mathbb{I} \neq \sum_n |\psi_n\rangle\langle\psi_n|$ the states do exhibit P-orthogonality:

$$P = \sum_n |P\psi_n\rangle\langle P\psi_n|, \quad P|_{\mathscr{F}_\Lambda} = \mathbb{I}. \tag{4.16}$$

As we can only ensure orthogonality on $\mathscr{F}_\Lambda$, a consistent spectral representation can only be constructed from the projectors $|P\psi_n\rangle\langle P\psi_n|$. The unique operator on $\mathscr{F}_\Lambda$ with spectrum $\{\omega_n\}_n$ and eigenvectors $\{|P\psi_n\rangle\}_n$ is

$$\Omega_{\mathscr{F}_\Lambda} = \sum_n \omega_n |P\psi_n\rangle\langle P\psi_n|. \tag{4.17}$$

The operator $\Omega_{\mathscr{F}_\Lambda}$ is the dual space analogue of Ω. The analogue of $R(\omega)$ is the *projected resolvent* $\mathcal{R}(\omega)$, the resolvent of $\Omega_{\mathscr{F}_\Lambda}$:

$$\mathcal{R}(\omega) = \sum_n \frac{|P\psi_n\rangle\langle P\psi_n|}{\omega_n - \omega}. \tag{4.18}$$

We emphasize that any resolvent of the form $\sum_n f(\omega_n)|\psi_n\rangle\langle\psi_n|$ will be in contradiction with the Hermiticity of H by the non-orthogonality of the $|\psi_n\rangle$ (*e.g.* its square is inconsistent with the spectral theorem for H).

As an aside, we mention that the eigenvalue problem associated with $\Omega_{\mathscr{F}_\Lambda}$ is by construction equivalent to $\boldsymbol{\Omega}^r \mathbf{c}_n = \omega_n \boldsymbol{P}^r \mathbf{c}_n$. This can be seen by writing $\boldsymbol{P}^r = \boldsymbol{Q}^{r\dagger}\boldsymbol{Q}^r$ and rewriting the equation as

$$\boldsymbol{\Omega}_{\mathscr{F}_\Lambda} \mathbf{u}_n = \omega_n \mathbf{u}_n, \tag{4.19}$$

where $\boldsymbol{\Omega}_{\mathscr{F}_\Lambda} = (\boldsymbol{Q}^r)^{-\dagger}\boldsymbol{\Omega}^r(\boldsymbol{Q}^r)^{-1}$ and $\mathbf{u}_n = \boldsymbol{Q}^r\mathbf{c}_n$. We find that $\boldsymbol{\Omega}_{\mathscr{F}_\Lambda}$ and $\Omega_{\mathscr{F}_\Lambda}$ are equivalent by considering the spectral resolution of $\boldsymbol{\Omega}_{\mathscr{F}_\Lambda}$:

$$(\boldsymbol{\Omega}_{\mathscr{F}_\Lambda})_{\mu\nu} = \sum_n \omega_n (\boldsymbol{Q}^r\mathbf{c}_n\mathbf{c}_n^\dagger\boldsymbol{Q}^{r\dagger})_{\mu\nu} = \langle\mu|\Omega_{\mathscr{F}_\Lambda}|\nu\rangle. \tag{4.20}$$

The linear response function of symmetric equation of motion theory is defined by substituting $R(\omega)$ with the projected resolvent $\mathcal{R}(\omega)$ in $\langle\langle A; V\rangle\rangle_\omega$. As it is possible to show that

$$\mathcal{R}(\omega) = \sum_{\mu\nu} |\mu\rangle(\boldsymbol{Q}^r(\boldsymbol{\Omega}^r - \omega\,\boldsymbol{P}^r)^{-1}\boldsymbol{Q}^{r\dagger})_{\mu\nu}\langle\nu|, \tag{4.21}$$

the response function may be written

$$\langle\langle A; V^\omega\rangle\rangle_\omega^{\text{seom}} = -\bar{\mathbf{a}}^\dagger \boldsymbol{P}^r(\boldsymbol{\Omega}^r - \omega\,\boldsymbol{P}^r)^{-1}\boldsymbol{P}^r\,\bar{\mathbf{v}}^\omega -$$
$$\bar{\mathbf{v}}^{\omega\dagger}\boldsymbol{P}^r(\boldsymbol{\Omega}^r + \omega\,\boldsymbol{P}^r)^{-1}\boldsymbol{P}^r\,\bar{\mathbf{a}}, \tag{4.22}$$

where $\overline{a}_\mu = \langle \mu | e^{-T} A\, e^T | \mathrm{R} \rangle$ and $\overline{v}_\mu^\omega = \langle \mu | e^{-T} V^\omega e^T | \mathrm{R} \rangle$. Transition elements are now obtained by taking the appropriate limits:

$$\Gamma_{0\to k}^A \, \Gamma_{k\to 0}^{V^{\omega_k}} = \overline{\mathbf{a}}^\dagger \boldsymbol{P}^r \mathbf{c}_k \, \mathbf{c}_k^\dagger \, \boldsymbol{P}^r \, \overline{\mathbf{v}}^{\omega_k}, \tag{4.23}$$

$$\Gamma_{0\to k}^{V^{-\omega_k}} \, \Gamma_{k\to 0}^A = -\overline{\mathbf{v}}^{-\omega_k \dagger} \boldsymbol{P}^r \mathbf{c}_k \, \mathbf{c}_k^\dagger \, \boldsymbol{P}^r \, \overline{\mathbf{a}}. \tag{4.24}$$

The reasoning that led to $\mathcal{R}(\omega)$ is predicated on P-orthogonality. Alternative approximations $\mathcal{H}$ of H will therefore also have identical expressions for the molecular properties and transition elements (see Chapter 5).

Before moving on, we should remark that our approach is not analogous to that of Stanton and Bartlett.[22] Equation of motion theory prescribes the pseudo-projectors $|\psi_n\rangle\langle\psi_n^\Lambda|$ to approximate the spectral resolution of $R(\omega)$. The $\langle\psi_0|$ in Equation (4.14) is moreover substituted with $\langle\psi_0^\Lambda|$. The sense of this approach can be understood by observing that

$$\left(\sum_n |\psi_n^\Lambda\rangle\langle\psi_n^\Lambda| \right)^2 \neq \sum_n |\psi_n^\Lambda\rangle\langle\psi_n^\Lambda|, \quad \left(\sum_n |\psi_n\rangle\langle\psi_n^\Lambda| \right)^2 = \sum_n |\psi_n\rangle\langle\psi_n^\Lambda|. \tag{4.25}$$

Recall that projection operators P are defined by $P^2 = P$ and $P = P^{\dagger}$.[67] In the former case, $P = P^\dagger$ but $P^2 \neq P$; the situation is reversed in the latter. In our Hermitian theory we can ensure both properties simultaneously by constructing the resolvent's spectral resolution from the true projection operators $|P\psi_n\rangle\langle P\psi_n|$. This moreover allows us to leave the $\langle\psi_0|$ in Equation (4.14) as it is and place all approximations in the projected resolvent $\mathcal{R}(\omega)$.

4.1.5. The reduced symmetric equation of motion equation

The symmetric coupled cluster state is always a solution of the full equation of motion equation. As a consequence, it is of interest to derive an equation for the unknown excited states. We emphasize that for other choices of T (*e.g.* the traditional cluster amplitudes) the full equation of motion equation must be consulted.

We begin by writing $\boldsymbol{\Omega}^r$ and $\boldsymbol{P}^r$ in block form, separating the reference contributions from the rest. The blocks are made up of objects we encountered in symmetric coupled cluster theory:

$$\boldsymbol{P}^r = \begin{pmatrix} N & \mathbf{q}^\dagger \mathbf{Q} \\ \mathbf{Q}^\dagger \mathbf{q} & \mathbf{Q}^\dagger \mathbf{Q} \end{pmatrix}, \quad \boldsymbol{A}^r = \begin{pmatrix} \alpha & \boldsymbol{\eta}^\dagger \\ \boldsymbol{\xi} & \mathbf{A} \end{pmatrix}. \tag{4.26}$$

By multiplying $\boldsymbol{P}^r$ and $\boldsymbol{A}^r$, adding the adjoint and dividing by two, we find that every reference contribution vanishes. That is to say,

$$\boldsymbol{\Omega}^r = \frac{1}{2} \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \mathbf{Q}^\dagger \mathbf{Q}\, \mathbf{A} + \mathbf{Q}^\dagger \mathbf{q}\, \boldsymbol{\eta}^\dagger + \text{h.c.} \end{pmatrix} \equiv \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Omega} \end{pmatrix}. \tag{4.27}$$

This result is by construction: it is implied by stationarity of the ground state, *i.e.* the symmetric amplitude equations $0 = \partial\langle H\rangle_{\text{scc}}/\partial t_\mu = \partial\mathcal{H}_{00}/\partial t_\mu = \mathcal{H}_{0\mu}$.

As anticipated, the symmetric coupled cluster state is a solution with excitation energy equal to zero: $\mathbf{\Omega}^r \mathbf{c}_0 = \mathbf{0}$ for $\mathbf{c}_0 = (1\ \mathbf{0})^\dagger$. By virtue of their orthogonality to $\mathbf{c}_0$, eigenvectors $\mathbf{c}_n$ with $\omega_n \neq 0$ must satisfy

$$\mathbf{c}_0^\dagger \boldsymbol{P}^r \mathbf{c}_n = (1\ \mathbf{0}) \begin{pmatrix} N & \mathbf{q}^\dagger \mathbf{Q} \\ \mathbf{Q}^\dagger \mathbf{q} & \mathbf{Q}^\dagger \mathbf{Q} \end{pmatrix} \begin{pmatrix} c \\ \mathbf{c} \end{pmatrix} = 0 \implies c = -N^{-1} \mathbf{q}^\dagger \mathbf{Q} \, \mathbf{c}. \quad (4.28)$$

Thus

$$\boldsymbol{P}^r \mathbf{c}_n = \begin{pmatrix} Nc + \mathbf{q}^\dagger \mathbf{Q} \, \mathbf{c} \\ \mathbf{Q}^\dagger \mathbf{q} \, c + \mathbf{Q}^\dagger \mathbf{Q} \, \mathbf{c} \end{pmatrix} = \begin{pmatrix} 0 \\ \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q} \, \mathbf{c} \end{pmatrix}, \quad (4.29)$$

where $\mathbf{M} = \mathbf{I} - N^{-1} \mathbf{q} \, \mathbf{q}^\dagger$. To further simplify $\mathbf{\Omega}$, we substitute for $\boldsymbol{\eta}^\dagger$ by using the symmetric amplitude equations. One may then show that

$$\mathbf{\Omega} = \mathbf{Q}^\dagger \mathbf{Q} \, \mathbf{A} + \mathbf{Q}^\dagger \mathbf{q} \, \boldsymbol{\eta}^\dagger + \text{h.c.} = \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q} \, \mathbf{A} - \mathbf{Q}^\dagger \mathbf{K} \, \mathbf{Q} + \text{h.c.}, \quad (4.30)$$

where $\mathbf{K} = N^{-1}(\alpha^* \mathbf{q} \, \mathbf{q}^\dagger + \mathbf{q} \, \boldsymbol{\xi}^\dagger \mathbf{Q}^\dagger)$. By a few straight-forward though slightly tedious manipulations we arrive at the attractive result

$$\mathbf{Q}^\dagger (\mathbf{K} + \mathbf{K}^\dagger) \mathbf{Q} = (\overline{\boldsymbol{\tau}} \, \boldsymbol{\xi}^\dagger) \, \mathbf{Q}^\dagger \mathbf{M} \mathbf{Q} + \mathbf{Q}^\dagger \mathbf{M} \mathbf{Q} \, (\boldsymbol{\xi} \, \overline{\boldsymbol{\tau}}^\dagger), \quad (4.31)$$

where we have made use of $2 \, \Re e \, \alpha = -N^{-1} (\boldsymbol{\xi}^\dagger \mathbf{Q}^\dagger \mathbf{q} + \mathbf{q}^\dagger \mathbf{Q}^\dagger \boldsymbol{\xi})$. Interestingly, in contrast to the inner product shift $\boldsymbol{\xi}^\dagger \overline{\boldsymbol{\tau}}$ in the ground state energy E, $\mathbf{K}$ causes a shift in the Jacobian $\mathbf{A}$ equal to the outer product of the vectors $\boldsymbol{\xi}$ and $\overline{\boldsymbol{\tau}}$. Writing $\boldsymbol{\mathcal{A}} = \mathbf{A} - \boldsymbol{\xi} \, \overline{\boldsymbol{\tau}}^\dagger$ we have in fact found that

$$\mathbf{\Omega} = \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q} \, \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^\dagger \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q}. \quad (4.32)$$

Finally, we pull all the pieces together, see Equations (4.27), (4.29), and (4.32). The reference block of the equation is trivially zero, while the excited block is the *reduced* symmetric equation of motion equation:

$$(\mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q} \, \boldsymbol{\mathcal{A}} + \boldsymbol{\mathcal{A}}^\dagger \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q}) \, \mathbf{c}_n = \omega_n \, \mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q} \, \mathbf{c}_n. \quad (4.33)$$

We have reinstated the index n by denoting the excited part of the eigenvector as $\mathbf{c} \equiv \mathbf{c}_n$, *i.e.* this vector does not contain the reference contribution c.

The reduced symmetric equation of motion equation also admits a basis of eigenvectors $\mathbf{c}_n$ with real eigenvalues ω_n. We prove this by showing that $\mathbf{M}$ is positive definite. The claim then follows by $\mathbf{Q}^\dagger \mathbf{M} \, \mathbf{Q}$'s positive definiteness and the theorems presented in Section 4.1.3. Let us write out $\mathbf{M}$ in detail:

$$\mathbf{M} = \mathbf{I} - \frac{\mathbf{q} \, \mathbf{q}^\dagger}{1 + \mathbf{q}^\dagger \mathbf{q}}. \quad (4.34)$$

Clearly, $\mathbf{M} \mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in \text{span} \{\mathbf{q}\}^\perp$, meaning that 1 is an $\dim(\mathbf{M}) - 1$ degenerate eigenvalue of $\mathbf{M}$. The last eigenvector is $\mathbf{q}$ itself:

$$\mathbf{M} \mathbf{q} = \left(\frac{1}{1 + \mathbf{q}^\dagger \mathbf{q}} \right) \mathbf{q}. \quad (4.35)$$

With every eigenvalue of $\mathbf{M}$ being greater than zero, $\mathbf{M}$ must be positive definite by the variational theorem, $\mathbf{x}^\dagger \mathbf{M} \mathbf{x} \geq \inf \sigma(\mathbf{M})$ for $\mathbf{x}^\dagger \mathbf{x} = 1$.[67] The positive definiteness of $\mathbf{Q}^\dagger \mathbf{M} \mathbf{Q}$ now follows from $\mathbf{x}^\dagger \mathbf{Q}^\dagger \mathbf{M} \mathbf{Q} \mathbf{x} = \mathbf{y}^\dagger \mathbf{M} \mathbf{y}$. Strict positivity is guaranteed by $\mathbf{Q}$'s invertibility, since it ensures that $\mathbf{y} = \mathbf{Q} \mathbf{x}$ vanishes if and only $\mathbf{x}$ does.[85]

Chapter 5

Concluding remarks

One of the major motivations for a *symmetric* formulation of coupled cluster theory is that it suffers from none of the issues belaboring non-hermitian theories at conical intersections, now understood to be both ubiquitous in organic molecules and essential to describe their photochemistry. A prominent example in this regard is the chemical reactivity of the rhodopsin protein, where a conical intersection is thought responsible for the main photochemical event in vision.[111] The novel formulation presented in this thesis, symmetric coupled cluster theory, offers a partial solution to the long-standing problem raised by Hättig[48] regarding coupled cluster theory's inability to describe intersections between excited states.

In particular, the ground and excited state energies of the symmetric theory are never complex, and the excited state eigenvalue problem possesses the hermitian symmetry claimed[48] necessary to predict conical intersections with the correct dimension. In molecular systems comprised of fewer than 50 electrons, the accuracy of the theory's ground state energies (SCCSD) are moreover found to rival those of the traditional approach (CCSD). These results are encouraging, though it remains to be seen whether similar accuracy is maintained for the excitation energies, molecular properties, and transition elements. We have no implementation of these quantities at the time of writing.

Large molecular systems cannot be described by the method, as its energy and other quantities do not scale correctly with system size. Our study indicates that the model is accurate only for systems with no more than about 50 electrons. This places many organic molecules of interest out of reach, though certain model systems may be treated. For instance, an extensively used model system of the chromophore of the rhodopsin protein is within reach, the penta-2,4-dieniminium (PSB3) cation.[40,44] This limitation is nevertheless severe, and a priority for future research is the formulation of alternative symmetric approaches that exhibit better scaling properties.

Our research group and collaborators have discovered that CCSD and CC3 predict *real* excitation energies at a conical intersection between two excited states of thymine.[112] We have moreover observed a conical intersection in H_2S with CCSD that has the correct dimension.[113] These developments have impelled us to re-examine Hättig's analysis, and in this regard we have already made some progress (see Section 2.3.2). Whether the failure

of coupled cluster theory is the exception or the rule is not known, and a detailed assessment of this is therefore planned for future work. Based on its outcome, we may prioritize between the development and use of traditional or symmetric coupled cluster methods in excited state dynamics.

thymine   penta-2,4-dieniminium (PSB3)
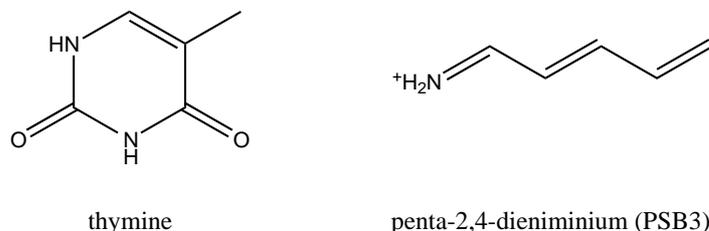
FIGURE 5.1: *Thymine and the penta-2,4-dieniminium (PSB3) cation.*

One of our long-term goals is the integration of coupled cluster methods in the *ab initio* molecular dynamics program developed by Martínez and coworkers.[45] To determine the dynamics this package requires excited state energy gradients and, whenever the coupling between electronic states becomes appreciable, the non-adiabatic coupling matrix elements. In order to implement these quantities, we may adopt established methods developed for traditional equation of motion coupled cluster theory.[114,115] Before we embark on this work, however, we will implement and assess the performance of the excitation energies, molecular properties, and transition elements of symmetric equation of motion theory.

Vital to continued work on symmetric coupled cluster theories is the proper consideration of alternative theories that could potentially offer superior scaling properties. We have recently considered the ground state expansion $|\psi\rangle = e^T \sum_{\mu \geq 0} c_\mu |\mu\rangle$, where both t_μ *and* c_μ are determined simultaneously. This is achieved by stationarity of the symmetric coupled cluster expectation value of H, giving the coupled set of equations

$$(\boldsymbol{\mathcal{H}}^r - E\,\boldsymbol{P}^r)\,\mathbf{c} = \mathbf{0},$$
$$\mathbf{c}^\dagger(\boldsymbol{\mathcal{H}}^r_\mu - E\,\boldsymbol{P}^r_\mu)\,\mathbf{c} = 0, \qquad (5.1)$$

where $\boldsymbol{\mathcal{H}}^r_\mu$ and $\boldsymbol{P}^r_\mu$ are non-hermitian matrices for each $\mu > 0$. Advantages are perhaps gained in the multi-configurational reference $\sum_{\mu \geq 0} c_\mu |\mu\rangle$, only endowed on the excited states in equation of motion theories based on the coupled cluster state.

Our framework moreover allows several generalizations of the expectation value $\langle H \rangle$. In the family of approximations $\langle H \rangle_{F,G} = \langle \psi | F(H) | \psi \rangle / \langle \psi | G | \psi \rangle$, one is of particular interest to us:

$$F(H) = PH + HP - PHP, \quad G = P. \qquad (5.2)$$

This is the unique choice that simultaneously satisfies the Hellmann-Feynman theorem while providing the correct block structure of $\mathbf{H}$ (see Figure 5.2). It

moreover offers a consistent operator approximation prescription,

$$F(H) = \hat{P}H, \quad G = \hat{P}\,\mathbb{I}, \tag{5.3}$$

where $\hat{P}\mathcal{O} = P\mathcal{O} + P\mathcal{O} - P\mathcal{O}P$. It is a consistent approach in the sense that $\hat{P}^2 = \hat{P}$; approximated operators remain invariant under $\hat{P}$. The amplitude equations and the energy may be derived by the procedure outlined for symmetric coupled cluster theory (see Section 3.1.4).

$$\mathbf{H} = \begin{pmatrix} & \text{S} & \text{D} & \text{T} & \text{Q} \\ & & 1 & & 1 \\ & & 1 & & 0 \end{pmatrix} \begin{matrix} \text{S} \\ \text{D} \\ \text{T} \\ \text{Q} \end{matrix}$$

FIGURE 5.2: *Matrix elements of the operator* $PH + HP - PHP$. *The relative sizes as compared to the exact* $\langle\varphi|H|\psi\rangle$ *are listed.*

There are two main advantages to employing P: it ensures finiteness and it provides optimal real approximations of $\langle H\rangle$ (see Sections 2.4.2 and 3.1.1). A caveat is that the resulting amplitude equations do *not* project the Schrödinger equation onto any subspace. Rigorous size-extensivity is lost as a consequence (see Section 3.3.7). Although entirely different approaches are of interest, it appears that finiteness and symmetry are not easily reconciled with Schrödinger equation projection. An improper scaling with system size places an upper limit on the treatable molecular size. Where this limit is met will however vary from model to model. Alternative theories might exhibit less severe limitations, and we remain optimistic that such symmetric coupled cluster theories exist.

Bibliography

1. T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular electronic-structure theory*, John Wiley & Sons, 2014.

2. F. Coester, Nucl. Phys. **7**, 421 (1958).

3. F. Coester and H. Kummel, Nucl. Phys. **17**, 477 (1960).

4. J. Čížek, J. Chem. Phys. **45**, 4256 (1966).

5. J. Čížek and J. Paldus, Int. J. Quant. Chem. **5**, 359 (1971).

6. G. D. Purvis III and R. J. Bartlett, J. Chem. Phys. **76**, 1910 (1982).

7. Y. S. Lee, S. A. Kucharski, and R. J. Bartlett, The Journal of chemical physics **81**, 5906 (1984).

8. R. J. Bartlett, Mol. Phys. **108**, 2905 (2010).

9. S. F. Boys, Proc. R. Soc. A **200**, 542 (1950).

10. C. Møller and M. S. Plesset, Phys. Rev. **46**, 618 (1934).

11. R. M. Dreizler and E. K. Gross, *Density functional theory: an approach to the quantum many-body problem*, Springer Science & Business Media, 2012.

12. R. J. Bartlett and M. Musiał, Rev. Mod. Phys. **79**, 291 (2007).

13. H. J. Monkhorst, Int. J. Quant. Chem. **12**, 421 (1977).

14. E. Dalgaard and H. J. Monkhorst, Phys. Rev. A **28**, 1217 (1983).

15. A. L. Fetter and J. D. Walecka, *Quantum theory of many-particle systems*, Courier Corporation, 2003.

16. H. Bruus and K. Flensberg, *Many-body quantum theory in condensed matter physics: an introduction*, Oxford University Press, 2004.

17. P. Langhoff, S. Epstein, and M. Karplus, Rev. Mod. Phys. **44**, 602 (1972).

18. J. Arponen, Ann. Phys. **151**, 311 (1983).

19. H. Koch and P. Jørgensen, J. Chem. Phys. **93**, 3333 (1990).

20. C. Hättig, O. Christiansen, and P. Jørgensen, Chem. Phys. Lett. **282**, 139 (1998).

21. T. Helgaker et al., Chem. Rev. **112**, 543 (2012).

22. J. F. Stanton and R. J. Bartlett, J. Chem. Phys. **98**, 7029 (1993).

23. P. Rozyczko and R. J. Bartlett, J. Chem. Phys. **107**, 10823 (1997).

24. P. B. Rozyczko and R. J. Bartlett, J. Chem. Phys. **108**, 7988 (1998).

25. C. Hättig et al., J. Chem. Phys. **109**, 3293 (1998).

26. P. B. Rozyczko and R. J. Bartlett, J. Chem. Phys. **109**, 9201 (1998).

27. J. F. Stanton and J. Gauss, J. Chem. Phys. **101**, 8938 (1994).

28. M. Nooijen and R. J. Bartlett, J. Chem. Phys. **102**, 3629 (1995).

29. A. I. Krylov, Annu. Rev. Phys. Chem. **59**, 433 (2008).

30. J. C. Slater, Phys. Rev. **81**, 385 (1951).

31. D. I. Lyakh, M. Musiał, V. F. Lotrich, and R. J. Bartlett, Chem. Rev. **112**, 182 (2012).

32. F. A. Evangelista, W. D. Allen, and H. F. Schaefer III, J. Chem. Phys. **125**, 154113 (2006).

33. P. G. Szalay, T. Muller, G. Gidofalvi, H. Lischka, and R. Shepard, Chem. Rev. **112**, 108 (2011).

34. D. R. Yarkony, Rev. Mod. Phys. **68**, 985 (1996).

35. G. A. Worth and L. S. Cederbaum, Annu. Rev. Phys. Chem. **55**, 127 (2004).

36. W. Domcke et al., *Conical intersections: electronic structure, dynamics and spectroscopy*, World Scientific, 2004.

37. M. Ben-Nun et al., Adv. Chem. Phys. **121**, 439 (2002).

38. E. Teller, J. Phys. Chem. **41**, 109 (1937).

39. D. R. Yarkony, Acc. Chem. Res. **31**, 511 (1998).

40. D. Tuna et al., J. Chem. Theory Comput. **11**, 5758 (2015).

41. S. Perun, , A. L. Sobolewski, and W. Domcke, J. Phys. Chem. A **110**, 13238 (2006).

42. S. Gozem et al., J. Chem. Theory Comput. **10**, 3074 (2014).

43. S. Gozem, A. I. Krylov, and M. Olivucci, J. Chem. Theory Comput. **9**, 284 (2013).

44. S. Gozem et al., J. Chem. Theory Comput. **9**, 4495 (2013).

45. M. Ben-Nun, J. Quenneville, and T. J. Martínez, J. Phys. Chem. A **104**, 5161 (2000).

46. B. G. Levine and T. J. Martínez, Annu. Rev. Phys. Chem. **58**, 613 (2007).

47. W. Domcke, D. R. Yarkony, and H. Köppel, *Conical intersections: theory, computation and experiment*, World Scientific, 2011.

48. C. Hättig, Structure optimizations for excited states with correlated second-order methods: CC2 and ADC(2), in *Response Theory and Molecular Properties (A Tribute to Jan Linderberg and Poul Jørgensen)*, edited by H. Jensen, volume 50 of *Adv. Quantum Chem.*, pages 37 – 60, Academic Press, 2005.

49. A. Köhn and A. Tajti, J. Chem. Phys. **127** (2007).

50. A. Dreuw and M. Wormit, Wiley Interdiscip. Rev. Comput. Mol. Sci. **5**, 82 (2015).

51. J. Schirmer and F. Mertins, Int. J. Quant. Chem. **58**, 329 (1996).

52. P. H. Harbach, M. Wormit, and A. Dreuw, J. Chem. Phys. **141**, 064113 (2014).

53. J. Schirmer, Phys. Rev. A **26**, 2395 (1982).

54. R. J. Bartlett, Int. J. Mol. Sci. **3**, 579 (2002).

55. O. Christiansen, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **243**, 409 (1995).

56. P. G. Szalay, M. Nooijen, and R. J. Bartlett, J. Chem. Phys. **103**, 281 (1995).

57. T. B. Pedersen and H. Koch, J. Chem. Phys. **106**, 8059 (1997).

58. S. Kvaal, Mol. Phys. **111**, 1100 (2013).

59. G. K.-L. Chan, M. Kállay, and J. Gauss, J. Chem. Phys. **121**, 6110 (2004).

60. R. P. Feynman, Phys. Rev. **56**, 340 (1939).

61. P. G. Szalay and R. J. Bartlett, Int. J. Quant. Chem. **44**, 85 (1992).

62. A. G. Taube and R. J. Bartlett, Int. J. Quant. Chem. **106**, 3393 (2006).

63. R. Yaris, J. Chem. Phys. **41**, 2419 (1964).

64. H. Koch et al., J. Chem. Phys. **100**, 4393 (1994).

65. R. Kobayashi, H. Koch, and P. Jørgensen, Chem. Phys. Lett. **219**, 30 (1994).

66. E. Schrödinger, Phys. Rev. **28**, 1049 (1926).

67. G. Teschl, *Mathematical methods in quantum mechanics*, American Mathematical Soc., 2014.

68. D. J. Griffiths, *Introduction to quantum mechanics*, Pearson Education, 2005.

69. P. W. Atkins and R. S. Friedman, *Molecular quantum mechanics*, Oxford University Press, 2011.

70. P. A. M. Dirac, Math. Proc. Cambridge **35**, 416 (1939).

71. L. E. Spence, A. J. Insel, and S. H. Friedberg, *Elementary linear algebra*, Pearson/Prentice Hall, 2008.

72. C. Gerry and P. Knight, *Introductory quantum optics*, Cambridge University Press, 2005.

73. E. A. Hylleraas and B. Undheim, Z. Phys. **65**, 759 (1930).

74. R. Shankar, *Principles of quantum mechanics*, Springer Science & Business Media, 2012.

75. E. Kreyszig, *Advanced engineering mathematics*, John Wiley & Sons, 1988.

76. H. Koch et al., J. Chem. Phys. **106**, 1808 (1997).

77. H. Koch, O. Christiansen, P. Jørgensen, and J. Olsen, Chem. Phys. Lett. **244**, 75 (1995).

78. J. D. Watts, J. Gauss, and R. J. Bartlett, J. Chem. Phys. **98**, 8718 (1993).

79. M. Nooijen, K. Shamasundar, and D. Mukherjee, Mol. Phys. **103**, 2277 (2005).

80. P. Atkins and J. De Paula, *Elements of physical chemistry*, Oxford University Press, 2013.

81. I. Shavitt and R. J. Bartlett, *Many-body methods in chemistry and physics: MBPT and coupled-cluster theory*, Cambridge University Press, 2009.

82. R. A. Adams, *Calculus: A complete course*, Addison-Wesley, 1995.

83. W. Domcke, D. Yarkony, and H. Köppel., *Electronic Structure, Dynamics and Spectroscopy*, volume 15, World Scientific, 2004.

84. T. Pacher, C. A. Mead, L. S. Cederbaum, and H. Köppel, J. Chem. Phys. **91**, 7057 (1989).

85. H. Anton, *Elementary linear algebra*, John Wiley & Sons, 2010.

86. C. D. Sherrill, A. I. Krylov, E. F. Byrd, and M. Head-Gordon, J. Chem. Phys. **109**, 4171 (1998).

87. T. B. Pedersen, H. Koch, and C. Hättig, J. Chem. Phys. **110**, 8318 (1999).

88. N. C. Handy, J. A. Pople, M. Head-Gordon, K. Raghavachari, and G. W. Trucks, Chem. Phys. Lett. **164**, 185 (1989).

89. R. J. Bartlett and J. Noga, Chem. Phys. Lett. **150**, 29 (1988).

90. J. B. Robinson and P. J. Knowles, J. Chem. Phys. **136**, 054114 (2012).

91. R. J. Bartlett, S. A. Kucharski, and J. Noga, Chem. Phys. Lett. **155**, 133 (1989).

92. W. Kutzelnigg, *Recent Progress in Coupled Cluster Methods: Theory and Applications*, chapter Unconventional Aspects of Coupled-Cluster Theory, pages 299–356, Springer Netherlands, Dordrecht, 2010.

93. R. Moszynski, P. S. Żuchowski, and B. Jeziorski, Collect. Czech. Chem. Commun. **70**, 1109 (2005).

94. H. Koch et al., J. Chem. Phys. **92**, 4924 (1990).

95. K. Aidas et al., WIREs Comput. Mol. Sci. **4**, 269 (2014).

96. P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).

97. P. Pulay, J. Comput. Chem. **3**, 556 (1982).

98. G. E. Scuseria, C. L. Janssen, and H. F. Schaefer, J. Chem. Phys. **89**, 7382 (1988).

99. T. H. Dunning, J. Chem. Phys. **90**, 1007 (1989).

100. J. Olsen et al., J. Chem. Phys. **104**, 8007 (1996).

101. Y. Shao et al., Mol. Phys. **113**, 184 (2015).

102. M. D. Hanwell et al., J. Cheminformatics **4**, 17 (2012).

103. E. F. Pettersen et al., J. Comput. Chem. **25**, 1605 (2004).

104. E. Kreyszig, *Advanced engineering mathematics*, John Wiley & Sons, 2010.

105. M. E. Khatib et al., J. Phys. Chem. A **118**, 6664 (2014).

106. B. Cooper and P. J. Knowles, J. Chem. Phys. **133** (2010).

107. F. A. Evangelista, J. Chem. Phys. **134**, 224102 (2011).

108. G. H. Golub and C. F. Van Loan, *Matrix computations*, volume 3, JHU Press, 2012.

109. Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst, *Templates for the solution of algebraic eigenvalue problems: a practical guide*, volume 11, Siam, 2000.

110. B. C. Hall, *Quantum theory for mathematicians*, Springer, 2013.

111. D. Polli et al., Nature **467**, 440 (2010).

112. T. J. A. Wolf, R. H. Myhre, S. Coriani, H. Koch, A. Battistoni, N. Berrah, P. Bucksbaum, R. Coffee, G. Coslovich, J.P. Cryan, R. Feifel, K. Gaffney, J. Grilj, T. J. Martinez, S. Myabe, S. P. Møller, M. Mucke, A. Natan, R. Obaid, T. Osipov, O. Plekan, A. Sage, R. Squibb, S. Wang, and M. Gühr, (To be submitted to Science and appear on arXiv.org) .

113. E. F. Kjønstad, R. H. Myhre, T. J. Martinez, and H. Koch, (To be submitted and appear on arXiv.org) .

114. T. Ichino, J. Gauss, and J. F. Stanton, J. Chem. Phys. **130**, 174105 (2009).

115. J. F. Stanton, J. Chem. Phys. **99**, 8840 (1993).

List of symbols

α	The number $\langle \mathrm{R} | A | \mathrm{R} \rangle$.
$\langle \mathcal{O} \rangle$	The expectation value of $\mathcal{O}$.
$\mathcal{O}^r$	A matrix representation of $\mathcal{O}$ in a reference-including basis, either $\{|\mathrm{R}\rangle, |\mu\rangle\}_\mu$ or $\{e^T |\mathrm{R}\rangle, e^T |\mu\rangle\}_\mu$ depending on $\mathcal{O}$.
δ	The energy shift of the symmetric formulation, see p. 35.
$\mathcal{M}, \mathcal{M}_i$	Determinantal bases of $\mathscr{F}$ and $\mathscr{F}_i$.
$\mathscr{F}$	The N-electron Fock space, see p. 6.
$\mathscr{F}_i$	A subspace of $\mathscr{F}$ comprised of determinants of excitation order less than or equal to i, see p. 7.
$\Gamma^{\mathcal{O}}_{n \to m}$	The transition element between eigenstates n and m for the operator $\mathcal{O}$, see p. 9.
$\mathscr{H}$	The Hilbert space of square integrable functions.
$\mathbb{I}$	The identity operator.
P	The projection operator onto $\mathscr{F}_\Lambda$.
$\mathbf{A}$	The coupled cluster Jacobian, A written in the basis $\{|\mu\rangle\}_\mu$.
$|\mu\rangle$	The determinant $|\mu\rangle = \tau_\mu |\mathrm{R}\rangle$.
$|\psi_n^\Lambda\rangle$	The dual of $|\psi_n\rangle$, see p. 12.
$|\psi_\Lambda\rangle, |\Psi_\Lambda\rangle$	The duals of $|\psi\rangle$ and $|\Psi\rangle$, see p. 11.
$|\mathrm{R}\rangle$	The Hartree-Fock determinant.
λ	The perturbation strength or an arbitrary eigenvalue.
$\bar{t}_\mu$	A multiplier, see p. 11.
ω, ω_i	An excitation energy or a frequency.
$\mathcal{O}$	An arbitrary operator.
φ_i	A spin orbital.

$\mathbf{r}_i, \mathbf{R}_I$	The spatial coordinates of electrons and nuclei.
Ψ, ψ	The time-dependent quantum state and the wavefunction.
$\langle\langle A; V^\omega \rangle\rangle_\omega$	The linear response functions, see p. 9.
$\mathcal{H}$	The symmetric coupled cluster approximation of H, see p. 26.
Ω	A shifted $\mathcal{H}$, see p. 44.
$\mathcal{R}(\omega)$	The projected resolvent, see p. 46.
$\sigma(\mathcal{O})$	The spectrum of $\mathcal{O}$, see p. 6.
$\mathscr{F}_\psi, \mathscr{F}_\Lambda$	The state and dual spaces, see p. 12.
$\overline{H}$	The similarity transformed Hamiltonian, see p. 11.
τ_μ	An excitation operator, see p. 7.
$\overline{\boldsymbol{\tau}}$	The symmetric analogue of $\overline{\mathbf{t}}$, see p. 24.
$\boldsymbol{\eta}, \boldsymbol{\xi}$	Vectors associated with the Jacobian A, see p. 24.
$\mathbf{q}, \mathbf{Q}$	Objects associated with P, see p. 24.
A	The Jacobian (see p. 11) or an arbitrary operator.
$a, a^\dagger$	Annihilation and creation operators.
E	The energy.
H	The Hamiltonian operator.
H_0	The reference Hamiltonian in perturbation theory.
M	The number of spin orbitals in the given spin orbital basis.
N	A normalization (see p. 24) or the number of electrons.
$R(\omega)$	The resolvent of H.
T	The cluster operator, see p. 10.
t_μ	The cluster amplitudes.
U	A propagator, see p. 9.
V	The perturbation operator.
V^ω	The Fourier transform of V.

Appendix A

Some results and theorems

Theorem A.1. *Let $|\psi\rangle = e^T \sum_{\mu \geq 0} c_\mu |\mu\rangle$. Suppose moreover that G is hermitian and $F(H)$ is a hermitian operator-valued function of $H = H_0 + \lambda V$ such that $\partial F(H)/\partial \lambda|_{\lambda=0} = 0$. Then the Hellmann-Feynman theorem is valid for*

$$\langle H \rangle_{F,G} = \frac{\langle \psi \, | \, F(H) \, | \, \psi \rangle}{\langle \psi \, | \, G \, | \, \psi \rangle} \tag{A.1}$$

provided $\partial \langle H \rangle_{F,G} / \partial t_\mu|_{\lambda=0} = 0$ and $\partial \langle H \rangle_{F,G} / \partial c_\mu|_{\lambda=0} = 0$.

Proof. By hermiticity, $\langle \psi | \mathcal{O} | \varphi \rangle = \langle \varphi | \mathcal{O} | \psi \rangle^*$ for $\mathcal{O} = F(H)$ and G. Let $|\psi_{\vartheta_\mu}\rangle = \partial |\psi\rangle / \partial \vartheta_\mu$, $\vartheta_\mu \in \{c_\mu, t_\mu\}$. Then $\partial \langle \psi | / \partial \vartheta_\mu^* = \langle \psi_{\vartheta_\mu} |$ and, by $\langle \psi_{\vartheta_\mu} | \mathcal{O} | \psi \rangle = \langle \psi | \mathcal{O} | \psi_{\vartheta_\mu} \rangle^*$, we obtain the needed symmetry relation:

$$\frac{\partial \langle H \rangle_{F,G}}{\partial \vartheta_\mu} = \left(\frac{\partial \langle H \rangle_{F,G}}{\partial \vartheta_\mu^*} \right)^*. \tag{A.2}$$

Clearly, moreover, $\partial \langle H \rangle_{F,G} / \partial \lambda|_{\lambda=0} = \langle V \rangle_{F,G}$. The theorem now follows by differentiation of $\langle H \rangle_{F,G}$ with respect to λ (see p. 22). $\qquad \square$

Theorem A.2. *Consider a subspace $\mathscr{F}_0 \subseteq \mathscr{F}$ and a vector $|\psi\rangle \in \mathscr{F}$. The orthogonal projection $P|\psi\rangle$ onto $\mathscr{F}_0$ is the closest vector (in terms of norm) to $|\psi\rangle$ in that subspace.*

Proof. Let $\{|\varphi_n\rangle\}_{n=1}^{N'}$ be an orthonormal basis for $\mathscr{F}_0$. Now extend this orthonormal basis $\{|\varphi_n\rangle\}_{n=1}^{N}$ such that it forms a basis for $\mathscr{F}$. Let us write $|\psi\rangle = \sum_{n=1}^{N} c_n |\varphi_n\rangle$. For $|\psi'\rangle = \sum_{n=1}^{N'} c_n' |\varphi_n\rangle \in \mathscr{F}_0$, we then have

$$|\vartheta\rangle = |\psi\rangle - |\psi'\rangle = \sum_{n=1}^{N'} (c_n - c_n') |\varphi_n\rangle + \sum_{m=N'+1}^{N} c_n |\varphi_n\rangle \tag{A.3}$$

and so

$$\langle \vartheta \, | \, \vartheta \rangle = \sum_{n=1}^{N'} |c_n - c_n'|^2 + \sum_{m=N'+1}^{N} |c_n|^2. \tag{A.4}$$

This norm is clearly at the minimum when $c_n' = c_n$ for $n = 1, 2, \ldots, N'$, which is only the case if $|\psi'\rangle = P|\psi\rangle$. $\qquad \square$

Appendix B

Ammonia cluster input geometries

TABLE B.1: *Cartesian coordinates of the $(NH_3)_2$ system. Distances are listed in units of ångström.*

	x	y	z
N	0.0000000000	0.0000000000	0.0000000000
N	1.2395729503	0.2742076703	-2.7181370559
H	1.0200000000	0.0000000000	0.0000000000
H	-0.3400000000	-0.7192700000	0.6383200000
H	-0.3400000000	-0.1931700000	-0.9420600000
H	2.2595729503	0.2742076703	-2.7181370559
H	0.8995729503	1.1310076703	-2.2814570559
H	0.8995629503	-0.5323723297	-2.1944670559

TABLE B.2: *Cartesian coordinates of the $(NH_3)_3$ system. Distances are listed in units of ångström. The central nitrogen is denoted by an '∗'.*

	x	y	z
N∗	-1.7519800000	0.4309000000	3.9322700000
N	0.7024400000	0.4063100000	-1.5125600000
N	-0.5371300000	0.1321100000	1.2055800000
H	-2.0919800000	0.9501300000	3.1228300000
H	-2.0919800000	-0.5297200000	3.8873300000
H	-0.7319800000	0.4309000000	3.9322700000
H	0.3624300000	-0.4002700000	-0.9888900000
H	0.3624400000	1.2631100000	-1.0758800000
H	1.7224400000	0.4063100000	-1.5125600000
H	-0.8771300000	-0.0610600000	0.2635200000
H	-0.8771300000	-0.5871600000	1.8439000000
H	0.4828700000	0.1321100000	1.2055800000

TABLE B.3: *Cartesian coordinates of the $(NH_3)_4$ system. Distances are listed in units of ångström. The central nitrogen is denoted by an '∗'.*

	x	y	z
N∗	-0.8925500000	-2.8269400000	0.8625100000
N	-1.7519800000	0.4309000000	3.9322700000
N	0.7024400000	0.4063100000	-1.5125600000
N	-0.5371300000	0.1321100000	1.2055800000
H	-1.2325500000	-3.3463100000	1.6718700000
H	-1.2325400000	-1.8663300000	0.9076300000
H	0.1274500000	-2.8269400000	0.8625100000
H	-2.0919800000	0.9501300000	3.1228300000
H	-2.0919800000	-0.5297200000	3.8873300000
H	-0.7319800000	0.4309000000	3.9322700000
H	0.3624300000	-0.4002700000	-0.9888900000
H	0.3624400000	1.2631100000	-1.0758800000
H	1.7224400000	0.4063100000	-1.5125600000
H	-0.8771300000	-0.0610600000	0.2635200000
H	-0.8771300000	-0.5871600000	1.8439000000
H	0.4828700000	0.1321100000	1.2055800000

TABLE B.4: *Cartesian coordinates of the $(NH_3)_5$ system. Distances are listed in units of ångström. The central nitrogen is denoted by an '$*$'.*

	x	y	z
N$*$	0.4998700000	2.8612700000	1.8957300000
N	-0.8925500000	-2.8269400000	0.8625100000
N	-1.7519800000	0.4309000000	3.9322700000
N	0.7024400000	0.4063100000	-1.5125600000
N	-0.5371300000	0.1321100000	1.2055800000
H	0.1598700000	2.7882600000	0.9368400000
H	0.1598700000	2.0673500000	2.4384100000
H	1.5198700000	2.8612700000	1.8957300000
H	-1.2325500000	-3.3463100000	1.6718700000
H	-1.2325400000	-1.8663300000	0.9076300000
H	0.1274500000	-2.8269400000	0.8625100000
H	-2.0919800000	0.9501300000	3.1228300000
H	-2.0919800000	-0.5297200000	3.8873300000
H	-0.7319800000	0.4309000000	3.9322700000
H	0.3624300000	-0.4002700000	-0.9888900000
H	0.3624400000	1.2631100000	-1.0758800000
H	1.7224400000	0.4063100000	-1.5125600000
H	-0.8771300000	-0.0610600000	0.2635200000
H	-0.8771300000	-0.5871600000	1.8439000000
H	0.4828700000	0.1321100000	1.2055800000