ESAIM: M2AN 47 (2013) 421–447
DOI: 10.1051/m2an/2012035

**ESAIM: Mathematical Modelling and Numerical Analysis**
www.esaim-m2an.org

# THE CONTINUOUS COUPLED CLUSTER FORMULATION FOR THE ELECTRONIC SCHRÖDINGER EQUATION *

Thorsten Rohwedder[1]

**Abstract.** Nowadays, the Coupled Cluster (CC) method is the probably most widely used high precision method for the solution of the main equation of electronic structure calculation, the *stationary electronic Schrödinger equation*. Traditionally, the equations of CC are formulated as a nonlinear approximation of a Galerkin solution of the electronic Schrödinger equation, *i.e.* within a given discrete subspace. Unfortunately, this concept prohibits the direct application of concepts of nonlinear numerical analysis to obtain *e.g.* existence and uniqueness results or estimates on the convergence of discrete solutions to the full solution. Here, this shortcoming is approached by showing that based on the choice of an $N$-dimensional reference subspace $R$ of $H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$, the original, continuous electronic Schrödinger equation can be reformulated equivalently as a root equation for an infinite-dimensional nonlinear Coupled Cluster operator. The canonical projected CC equations may then be understood as discretizations of this operator. As the main step, continuity properties of the cluster operator $S$ and its adjoint $S^\dagger$ as mappings on the antisymmetric energy space $\mathbb{H}^1$ are established.

**Mathematics Subject Classification.** 65Z05, 81-08, 70-08.

Received April 13, 2011. Revised July 3, 2012.
Published online January 11, 2013.

## 1. Introduction

The Coupled Cluster (CC) approach was derived around 1960 in the field of atomic physics [13, 14, 28, 45], and later introduced in the context of quantum chemistry (see [12]). It is today the probably most widely applied tool in the calculation of ground state solutions of the stationary $N$-electron Schrödinger equation when high-accuracy results are demanded. In the variant of the CCSD(T) method [34], which can be applied to small to medium-sized molecules with reasonable computational effort, CC often provides results which are within the error bars of corresponding practical experiments [30]. CCSD(T) is therefore often referred to as the "golden standard of quantum chemistry".

The ground state problem for the electronic Schrödinger equation, for the numerical treatment of which the CC method is used, governs the physical behaviour of $N$ electrons in the Coulomb field of a fixed set of nuclei, see [23, 42, 48] for some main results. To admit for a sensible discretization and a mathematically sound

*Keywords and phrases.* Quantum chemistry, electronic Schrödinger equation, coupled cluster method, numerical analysis, nonlinear operator equation.

* *This work has been supported by the DFG SPP programs 1145 and 1324.*

[1] Sekretariat MA 5-3, Institut für Mathematik, TU Berlin, Straße des 17. Juni 136, 10623 Berlin, Germany.
rohwedde@math.tu-berlin.de

algorithmic treatment, it is in the context of numerical analysis best phrased as a weak operator eigenproblem for an eigenfunction $\underline{\Psi}$ describing the electronic ground state [48], *i.e.* "Find $\underline{\Psi} \in \mathbb{H}^1$ and $E \in \mathbb{R}$ such that

$$\langle \Phi,\ H\underline{\Psi}\rangle \quad = \quad E\langle \Phi,\ \underline{\Psi}\rangle \qquad\qquad \text{for all } \ \Phi \in \mathbb{H}^1, \tag{1.1}$$

and such that $E$ is the lowest eigenvalue of $H$". In this, the solution space $\mathbb{H}^1$ is a suitable energy (Sobolev) space consisting of antisymmetric functions, and the operator $H : \mathbb{H}^1 \to \mathbb{H}^{-1}$ is the weak $N$-electron Hamiltonian, mapping to the dual space of $\mathbb{H}^1$ (see Sect. 2). To treat the Schrödinger equation (1.1) in the way the CC method is canonically used (see *e.g.* the quantum chemical standard work [21]), three steps are taken:

(a) Galerkin discretization of (1.1): restriction to a discrete subspace $\mathbb{H}^1_D$ gives a (usually extremely high-dimensional) discrete eigenvalue problem for a function $\underline{\Psi}_D \in \mathbb{H}^1_D$,

$$\langle \Phi_D,\ H\underline{\Psi}_D\rangle \ = \ E_D\langle \Phi_D,\ \underline{\Psi}_D\rangle \qquad \text{for all } \ \Phi_D \in \mathbb{H}^1_D. \tag{1.2}$$

By quantum chemists, $\underline{\Psi}_D$ is called the "full Configuration Interaction (full CI) solution" of the discrete system (1.2).

(b) In a second step, the full-CI equation is equivalently re-parametrized by an exponential ansatz as follows: from a preliminary Hartree-Fock calculation (see *e.g.* [21]), one has an antisymmetrized rank-1-approximation $\Psi_0$ at hand (a *Slater determinant*, see Sect. 2), being in most cases a rather good approximation to the sought solution $\underline{\Psi}_D$. $\underline{\Psi}_D$ is then written as a so-called excitation of the reference solution $\Psi_0$,

$$\underline{\Psi}_D \quad = \quad (I + S)\Psi_0,$$

in which $S$ is a linear operator, namely a so-called *cluster operator* that maps the reference $\Psi_0$ to the sought correction $\Psi^* = \underline{\Psi}_D - \Psi_0$ (also *cf.* Sect. 2 for the exact definition). $\Psi_0$ fixed, solution of (1.2) is thus equivalent to the computation of a cluster operator $S$ such that

$$\langle \Phi_D,\ H(I + S)\Psi_0\rangle \ = \ E_D\langle \Phi_D,\ (I + S)\Psi_0\rangle \qquad \text{for all } \ \Phi_D \in \mathbb{H}^1_D. \tag{1.3}$$

By standard matrix algebra (see *e.g.* [40, 45]), every cluster operator of the form $I + S$ can also be expressed as the exponential of a cluster operator $T$, so that (1.3) can in a second step be rephrased as determination of $T$ such that

$$\langle \Phi_D, He^T\Psi_0\rangle \ = \ E_D\langle \Phi_D,\ e^T\underline{\Psi}_D\rangle \qquad \text{for all } \ \Phi_D \in \mathbb{H}^1_D, \tag{1.4}$$

or alternatively, because $e^{-T}$ is invertible, as the solution of

$$\langle \Phi_D, e^{-T}He^T\Psi_0\rangle \ = \ E_D\langle \Phi_D,\ \Psi_0\rangle \qquad \text{for all } \ \Phi_D \in \mathbb{H}^1_D. \tag{1.5}$$

for $T$. These are the nonlinear "full-CC" equations (1.5) which are equivalent to the "full-CI"-formulation (1.2) on the space $\mathbb{H}^1_D$, and which define a nonlinear root equation for a coefficient vector of so-called *cluster amplitudes* $(t_\alpha)_{\alpha\in\mathcal{M}}$ determining $T$. In contrast to the terms occurring in (1.4), (1.5) can be evaluated exactly [16, 21, 36] and is therefore the formulation almost exclusively used in practice.

(c) In a final step, only certain of the amplitudes $t_\alpha$ determining $T$ are used in the calculation. This corresponds to a further reduction of the test space $\mathbb{H}^1_D$ to a subspace $\mathbb{H}^1_d$, usually pushing practically relevant problems into the range of computability. The result is a reduced set of CC equations

$$\langle \Phi_d, e^{-T}He^T\Psi_0\rangle \ = \ E_d\langle \Phi_d,\ \Psi_0\rangle \qquad\qquad \text{for all } \ \Phi_d \in \mathbb{H}^1_d. \tag{1.6}$$

The selection criteria for basis functions included in the calculation normally base on the so-called "excitation level" of the basis functions, leading then *e.g.* to the often used Coupled Cluster Singles Doubles (CCSD)

equations. In practice, the resulting equations are then evaluated with the aid of the Second Quantization formalism (see [16] for a comprehensible treatment) and then usually solved by Newton-type methods [21], often enhanced by the DIIS acceleration method [37].

In contrast to (1.5), the equations (1.6) are no longer equivalent to the CI (Galerkin) discretization of (1.1) on $\mathbb{H}^1_d$, but preferable over the CI method due to various favourable properties: The CC method enjoys a wide range of applicability in a black-box style and converges quickly and systematically to the full-CI energy limit $E_D$ when applied to relatively well-behaved systems as typically C-H-chains, rings, alcohols, ketones and amino acids are. It also usually outperforms the correspondingly truncated CI method, see *e.g.* [15, 26]. As another important feature truncated CC has, in contrast to truncated CI methods, the property of being size-consistent [2,4,32,40], making CC the tool of choice when describing reaction mechanisms. For a review on Coupled Cluster theory, the reader is referred to [3, 29] and the abundance of references given therein, as well as to the article [7] for a broader scope on the applications in physics; for some recent developments, see [5, 10, 27, 31] as well as the references given in Section 2.

In spite of the CC method's practical utility and popularity, theoretical results from the mathematical point of view are rather scarce. Only recently a first approach has been undertaken in [40], where the approximation properties of the truncation step from the discrete full-CI equations (1.3) to the projected Coupled Cluster equation (1.6) was analyzed. Thus, the problems associated with the direct re-formulation of the original, infinite-dimensional problem (1.1) as an infinite-dimensional nonlinear Coupled Cluster method approached in this work are circumvented; the flipside of this proceeding is that the results do not allow for direct estimates with respect to the true solution $\underline{\Psi} \in \mathbb{H}^1$, and convergence to $\underline{\Psi}$ can only be proved under certain uniformity assumptions for the discrete equations. Also, the approach *a priori* excludes the analysis of methods where the size of the underlying one-particle basis is varied. The latter are of interest in the context of error estimation though, especially in view of the fact that convergence of different CC models towards the limit within the full CI-space usually is rather fast, while the convergence of the full-CI solutions $\underline{\Psi}_D \in \mathbb{H}^1_D$ to the continuous limit $\underline{\Psi} \in \mathbb{H}^1$ is often rather slow with respect to the size of the underlying one-particle basis set.

As a first step towards such statements and error estimates, it is the goal of the present work to show that under suitable assumptions, the electronic Schrödinger equation (1.1) can in a mathematically rigorous fashion be *equivalently* re-formulated as Coupled Cluster equations in a coefficient space reflecting the continuous space $\mathbb{H}^1$. The resulting method will be termed "the continuous Coupled Cluster method", expressing the task of solving the eigenvalue equation (1.1) as that of determining a suitably defined cluster operator $T$ and an $E \in \mathbb{R}$ such that there holds

$$\langle \Phi, e^{-T} H e^{T} \Psi_0 \rangle = E \langle \Phi, \Psi_0 \rangle \qquad \text{for all } \ \Phi \in \mathbb{H}^1 \tag{1.7}$$

(with fixed reference $\Psi_0$). The set of all such $E$ then equals the set of all eigenvalues of (1.1), and if there are any eigenvalues below the essential spectrum of $H$, the smallest such $E$ thus gives the ground state energy of the system as above. The generalization of the canonical finite-dimensional CC formulation of the CI problem (1.2) to a continuous CC formulation (1.7) of the original problem (1.1) consists in the below three steps (i) to (iii) that will be taken care of in the following Sections 2 to 5:

(i) The formalism of cluster operators, canonically defined by their action on a fixed finite dimensional tensor tensor basis set, has to be adapted to the infinite dimensional space appropriately. This will be done in Section 2: To set the mathematical stage, we review the necessary parts of the ample mathematical background that underlies the electronic Schrödinger equation [22, 35, 44, 48]. We introduce annihilation and creation operators from Second Quantization, which are used to give the formal definition of cluster operators $S$ in infinite dimensional spaces. We then re-phrase the Configuration Interaction equations in terms of cluster operators and prove a one-to-one correspondence between corrections $\Psi^* \in \{\Psi_0\}^\perp$ and cluster operators $S \in B(\mathbb{L}_2, \mathbb{L}_2)$ in Theorem 2.7.

(ii) The critical point in establishing a functional-analytically sound generalization of the CC approach to infinite dimensional spaces consists in proving a one-to-one correspondence between the corrections $\Psi^* \in$

$\mathbb{H}^1 \cap \{\Psi_0\}^\perp$ contained in the energy space $\mathbb{H}^1$ and the cluster operators $S$ for which $S \in B(\mathbb{H}^1, \mathbb{H}^1)$. As well, the $\mathbb{H}^1$-continuity of their $L_2$-adjoints has to be established. The according statement are formulated in Theorem 4.1, the first main result of this work. To prove Theorem 4.1, there are to our knowledge no suitable concepts available in the literature so far. The idea of the unfortunately rather technical proof given in this work bases on representing a suitable norm on $\mathbb{H}^1$ by projection on orthogonal basis sets; the derivation of this representation is subject of Section 3. In Section 4, this is combined with certain consequences of the nilpotency properties of annihilation and creation operators, allowing to reduce the Proof of Theorem 4.1 to finite-dimensional, uniformly bounded $\ell_p$-estimates similar to the ones used [40].

(iii) The above continuity properties verified, application of well-known Banach algebra theory can be used to supply the remaining ingredients for formulation of the continuous CC equations and the continuous CC function $f$. This step is taken in Section 5. The second main result of this work is Theorem 5.3, where equivalence of the electronic Schrödinger equation to the full-CI equation is proved.

In a follow-up publication [38], we will harvest the continuous CC formulation to directly derive from it existence and uniqueness results for the continuous and discrete equations. We will show that the $\mathbb{H}^1$-error of the solutions of the discrete equations are bounded up to a constant by the $\mathbb{H}^1$-distance of the used approximation space to the full solution (*i.e.* we will give what in the context of Galerkin theory often is called *quasi-optimality estimates*). Also, we will obtain error estimators for the energies calculated by discrete CC, and (positive) statements on their convergence towards the real energy $E$. The corresponding analysis will also underpin the importance of particular constants (as the quality of the reference determinant $\Psi_0$ and spectral gaps of the Hamiltonian) for the practical convergence behaviour of the Coupled Cluster method.

## 2. Setting for the continuous CC equations

This section supplies the necessary building blocks for establishing the continuous Coupled Cluster method in the following sections. Based on the choice of a one-particle reference space $R \subseteq H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$, we start by re-phrasing the electronic Schrödinger equation (1.1) in its intermediate normalization formulation. This perturbational formulation is sometimes also termed the complete Configuration Interaction formulation (as in contrast to the above discrete full-CI formulation (1.2)) and is in its discrete version commonly used for the formulation of post-Hartree-Fock methods. We then introduce some concepts from Second Quantization; in particular, we define the excitation operators that are central to the formulation of the Coupled Cluster theory and compile some of their properties. Based on this, we write out the CI ansatz in terms of cluster operators and emphasize the problems that we will have to deal with in the following sections when making the transition to the exponential parametrization of the Coupled Cluster ansatz.

**The complete-CI formulation of the electronic Schrödinger equation.** The solution space $\mathbb{H}^1$ on which the $N$-particle electronic Schrödinger equation is formulated combines two requirements on the solution $\underline{\Psi}$: The first is that it be subject to the *Pauli principle*, according to which the wave function has to be antisymmetric (*i.e.* sign-changing) under every permutation of two non-identical particle coordinates $(x_i, s_i), (x_j, s_j) \in \mathbb{R}^3 \times \{\pm\frac{1}{2}\}$. This condition means that solutions have to be contained in the space[2]

$$\mathbb{L}^2 \;:=\; \mathbb{L}^2_N \;:=\; \bigwedge_{i=1}^{N} L^2\left(\mathbb{R}^3 \times \left\{\pm\frac{1}{2}\right\}\right) \;\subseteq\; L^2\left(\left(\mathbb{R}^3 \times \left\{\pm\frac{1}{2}\right\}\right)^N\right), \tag{2.1}$$

formed by all antisymmetric functions in $L^2((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N)$. The usual $L^2$-inner product and norm will be denoted by $\langle\cdot,\cdot\rangle$ resp. $\|\cdot\|$ in the following. The second requirement on the solution is that it be contained

[2]Note that when denoting the space $\mathbb{L}^2_N$ and later also $\mathbb{H}^1_N$, we will for brevity suppress the index $N$ in any context where the electron number is fixed.

in the Sobolev space $H^1\big((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N\big)$ of finite kinetic energy, a Hilbert space [39] with the inner product induced by the non-negative operator

$$N - \Delta \quad = \quad \sum_{i=1}^{N} (1 - \Delta)_{x_i}. \tag{2.2}$$

Herein, $\Delta$ is the $3N$-dimensional Laplace operator; in the right formula, $N - \Delta$ is expressed as a sum over $N$ three-dimensional operators $(1 - \Delta)_{x_i}$, acting on the variable $x_i$ of an $N$-particle function. The according inner product is given explicitly by

$$\langle \Phi, \Psi \rangle_1 \quad := \quad N\langle \Phi, \Psi \rangle + \langle \nabla\Phi, \nabla\Psi \rangle \quad := \quad N\langle \Phi, \Psi \rangle + \sum_{i=1}^{N} \langle \nabla_{x_i}\Phi, \nabla_{x_i}\Psi \rangle. \tag{2.3}$$

The energy space for the $N$-electron Schrödinger equation is now defined as

$$\mathbb{H}^1 \quad := \quad \mathbb{H}^1_N \quad := \quad \mathbb{L}^2_N \cap H^1\left(\left(\mathbb{R}^3 \times \left\{\pm\frac{1}{2}\right\}\right)^N\right). \tag{2.4}$$

Using the above inner product $\langle \cdot, \cdot\cdot \rangle_1$, $\mathbb{H}^1$ is also a Hilbert space. On $\mathbb{H}^1$, we are looking for solutions of the electronic Schrödinger equation. Its weak formulation, a convenient starting point for numerical treatment, is formulated in terms of the bounded, coercive [48] bilinear form $h : \mathbb{H}^1 \times \mathbb{H}^1 \to \mathbb{R}$ on the energy space $\mathbb{H}^1$, induced by the strong Hamiltonian $\widehat{H} : \mathbb{H}^2 \to \mathbb{L}^2$ [17, 22, 35, 42, 47],

$$h(\Psi, \Psi') := \langle \nabla\Psi, \nabla\Psi' \rangle \; + \; \left\langle \Psi, \left( \sum_{i=1}^{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \frac{1}{|x_i - x_j|} \; - \; \sum_{i=1}^{N} \sum_{k=1}^{M} \frac{Z_k}{|x_i - R_k|} \right) \Psi' \right\rangle. \tag{2.5}$$

In this, the constants $Z_k \in \mathbb{N}$, $R_k \in \mathbb{R}^3$ are the charges and positions of the $M$ fixed nuclei. The solutions of the weak eigenvalue equation

$$h(\underline{\Psi}, \cdot) \quad = \quad E\langle \underline{\Psi}, \cdot \rangle \quad \text{in} \quad \mathbb{H}^{-1}, \tag{2.6}$$

correspond to the eigenfunctions of the classical, unbounded Hamiltonian $\widehat{H} : \mathbb{H}^2 \to \mathbb{L}^2$ [48]. By standard functional analysis, (2.6) can be restated as an operator eigenvalue equation for a weak Hamiltonian $H : \mathbb{H}^1 \to \mathbb{H}^{-1}$,

$$\langle H\underline{\Psi}, \cdot \rangle \quad := \quad h(\underline{\Psi}, \cdot) \quad = \quad E\langle \underline{\Psi}, \cdot \rangle \quad \text{in} \quad \mathbb{H}^{-1}, \tag{2.7}$$

leading to the equation (1.1) formulated at the beginning of this work.

The reformulation of (2.7) in terms of the linear CI ansatz (*cf.* the introduction) and also in terms of the due continuous CC formulation are perturbational in the sense that they assume that an approximate, preliminarily calculated reference solution is given. In canonical, single-reference Coupled Cluster theory, this reference solution is an antisymmetrized rank-one approximation $\Psi_0$, mostly given by the Hartree–Fock solution [11, 21] in practice. Our formalism developed here will also be based on such a simple reference solution, characterizable equivalently by an $N$-dimensional one-particle reference space $R \subseteq H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$. The according construction is performed in the following fundamental definitions.

**Definition 2.1** (one-particle functions, antisymm. mapping, reference determinant)**.**

(i) Let $H^1 := H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$ denote the so-called space of one-particle functions, and fix an $N$-dimensional subspace $R$ of $H^1$. $R$ is called the occupied space of the ansatz, its $L^2$-orthogonal complement $R^\perp$ the virtual space.

(ii) By $\mathcal{Q} : L^2((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^N) \to \mathbb{L}^2$, we denote the $N$-particle antisymmetrization mapping, defined by its action on functions $\Psi = \Psi((x_1, s_1), \ldots, (x_N, s_N)) \in \mathbb{H}^1$ *via*

$$\mathcal{Q}\Psi = \frac{1}{\sqrt{N!}} \sum_{\pi \in S(N)} (-1)^{\mathrm{sgn}(\pi)} \Psi((x_{\pi(1)}, s_{\pi(1)}), \ldots, (x_{\pi(N)}, s_{\pi(N)})), \tag{2.8}$$

with the sum running over the permutational group $S(N)$ on $N$ elements operating on the indices of $\Psi$.[3]

(iii) Fixing an $L^2$-orthonormal basis

$$B_{\mathrm{occ}} \;:=\; \{\chi_1, \ldots, \chi_N\} \;\subseteq H^1 \tag{2.9}$$

of the space $R$ from *(i)*, the *reference determinant* $\Psi_0$ defined by $R$ is the normalized, antisymmetric function

$$\Psi_0 := \bigwedge_{I=1}^{N} \chi_I := \mathcal{Q}(\otimes_{I=1}^{N} \chi_I). \tag{2.10}$$

It is not hard to see that up to a phase factor, the reference $\Psi_0$ only depends on the chosen subspace $R$, not on the particular $B_{\mathrm{occ}}$. One mild restriction imposed on the choice of $R$ resp. $\Psi_0$ is the following.

**Assumption 2.2.** The reference $\Psi_0$ is not orthogonal to the sought ground state function,

$$\langle \underline{\Psi}, \Psi_0 \rangle \;\neq\; 0. \tag{2.11}$$

This assumption satisfied, equation (1.1) can now be formulated in terms of the CI ansatz: "Find $\underline{\Psi} = \Psi_0 + \Psi^* \in \mathbb{H}^1$, $\Psi^* \in \{\Psi_0\}^\perp$, such that

$$\langle \Phi, H(\Psi_0 + \Psi^*) \rangle \;=\; E\langle \Phi, \Psi_0 + \Psi^* \rangle \quad \text{for all} \quad \Phi \in \mathbb{H}^1, \quad \text{where} \quad \Psi^* \perp \Psi_0." \tag{2.12}$$

In this, the correction $\Psi^*$ is orthogonal to the reference in the $\mathbb{L}^2$-inner product, so that the condition

$$\langle \underline{\Psi}, \Psi_0 \rangle \;=\; 1 \tag{2.13}$$

is fulfilled, a condition often termed as *intermediate normalization* in quantum chemistry literature. The constraint (2.13) accounts for the fact that the solution of the eigenvalue equation (2.7) is always non-unique in the sense that it is fixed at most up to a constant factor; instead of the "usual" constraint that $\|\underline{\Psi}\| = 1$, (2.13) has the benefit that the set of all possible corrections $\Psi^*$ in (2.12) is now the *linear* space $\{\Psi_0\}^\perp \cap \mathbb{H}^1$.

**One-particle bases and Slater determinants.** We use the orthonormal basis $B_{\mathrm{occ}}$ of the reference space $R$ fixed above, and complement this by choosing an orthogonal basis $B_{\mathrm{virt}}$ of $R^\perp$,

$$B_{\mathrm{virt}} := \{\chi_A \mid \chi_A \in H^1,\ A \geq N+1\}, \qquad \text{span } B_{\mathrm{virt}} \;=\; R^\perp, \qquad B := B_{\mathrm{occ}} \cup B_{\mathrm{virt}}.$$

$B$ is a *one-particle basis*, a basis of $H^1(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$. Following the conventions of quantum chemistry literature, we will denote indices belonging to occupied orbitals $\chi_I \in B_{\mathrm{occ}}$ by letters

$$I, J, K, \ldots \in \mathrm{occ} := \{1, \ldots, N\}.$$

Virtual orbitals $\chi_A \in B_{\mathrm{virt}}$ are labelled by letters

$$A, B, C, \ldots \in \mathrm{virt} := \{M \in \mathbb{N} \mid M > N\},$$

Indices of unspecified orbitals $\chi_P$ are indicated by letters $P, Q, R, \ldots \in \mathbb{N}$.

[3]Note that $\mathcal{Q}$ is not the projector on the antisymmetric functions, but is designated for mapping the tensor product $\otimes_{I=1}^{N} \chi_{P_I}$ of $N$ orthonormal functions to a normalized, antisymmetric function as in (iii).

We now define for any set of $k$ indices $P_1, \ldots, P_k \in \mathbb{N}$, $k \in \mathbb{N}$, $P_i \neq P_j$ for $i \neq j$, the corresponding *$k$-particle Slater determinant* as

$$\Psi_\mu \;=\; \Psi[P_1, \ldots, P_k] := \mathcal{Q}_k\big(\otimes_{i=1}^k \chi_{P_i}\big) = \bigwedge_{i=1}^N \chi_{P_i}, \tag{2.14}$$

with $\mathcal{Q}_k$ denoting the $k$-particle antisymmetrisation mapping defined in (2.8). As in the above, Slater determinants will be labelled by an ordered multi-index $\mu = [P_1, \ldots, P_k]$, sometimes abbreviated by a Greek letter. For a Slater determinant $\Psi_\mu = \Psi[P_1, \ldots, P_k]$, we will also denote by

$$\mathrm{occ}(\mu) = \{P \in \{P_1, \ldots, P_k\} : P \in \mathrm{occ}\} \quad \text{and} \quad \mathrm{virt}(\mu) = \{P \in \{P_1, \ldots, P_k\} : P \in \mathrm{virt}\}$$

the indices of the occupied resp. virtual orbitals contained in $\mu$ and write $P \in \mu$ if $P \in \{P_1, \ldots, P_k\}$.

**Creation and annihilation operators.** The Coupled Cluster method utilizes a nonlinear parametrization of the space $\{\Psi_0\}^\perp \cap \mathbb{H}^1$ in terms of so-called excitation operators $X_\mu$, connected to the basis set introduced above in a natural way. The building blocks of such operators are so-called annihilation and creation operators borrowed the formalism of second quantization [6] that will now be introduced. In fact, any linear operator on $\mathbb{L}^2$ (in particular also the electronic Hamiltonian) may be written as a (possibly infinite) sum of polynomials of such creation and annihilation operators $a_P^\dagger, a_P$ (see [9] and also [46] for the equivalent concept of "matrix operators"). This allows to write the CI and CC formulations of the electronic Schrödinger equation (2.7) completely in terms of these operators, and the practical evaluation of the CC equations to be derived below essentially depends on the algebraic properties of these operators, see *e.g.* [16]. In contrast, our exposition here is quite short and only contains those properties that are needed later to establish the continuous Coupled Cluster method. For more details *cf.* [21, 43]; in particular, note that the proofs of the anticommutation laws and of the relations (2.21), (2.22) given below are analogous to those for the finite-dimensional case [21] and are therefore omitted.

We will in this paragraph utilize the antisymmetric, real valued space $\mathbb{L}^2 = \mathbb{L}^2_N$ from the last section for a varying number $N$ of electrons and therefore equip spaces, operators etc. under consideration with an index $N$ indicating the number of particles where needed. Because notations used are intuitive and only needed in this part, they will not be introduced at all length. From the next section on, the particle number $N$ will be fixed again; consequently, the indices will be omitted again. The (fermion) Fock space [18] is defined as

$$\mathbb{F} \;:=\; \bigoplus_{N=0}^\infty \mathbb{L}^2_N,$$

where $\bigoplus$ denotes the direct orthogonal sum of the $N$-electron Hilbert spaces $\mathbb{L}^2_N$. By writing $N$-electron state vectors $\Psi_N \in \mathbb{L}^2_N$ as $(\delta_{k,N}\Psi_N)_{k\in\mathbb{N}} = \;(0, 0, \ldots, 0, \Psi_N, 0, \ldots)$, we may embed $\mathbb{L}^2_N$ in $\mathbb{F}$ for any $N$. Note that the case $N = 0$ is also included in the above definition of the space $\mathbb{F}$. For this case, $\mathbb{L}^2_0$ is (by definition of the tensor product) the underlying field of the complex numbers. This is a one-dimensional vector space, thus containing up to a phase factor only one normalized vector called the *vacuum state* $|\rangle$. This state is in some sense the starting point for the formalism of second quantization, as any state vector may be created from it by the use of the creation operators introduced in the following.

**Definition 2.3** (creation and annihilation operators)**.** Let $f \in L^2(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$. The creation operator or creator $a_f^\dagger$ of $f$ is a linear mapping

$$a_f^\dagger : \mathbb{F} \to \mathbb{F} \tag{2.15}$$

defined by its action on each single space $\mathbb{L}^2_N$: For $N = 0$, we set

$$a_f^\dagger|\rangle \;=\; f \in \mathbb{L}^2_1;$$

for any $1 \leq N \in \mathbb{N}$, we define $a_f^\dagger|_{\mathbb{L}^2_N} : \mathbb{L}^2_N \to \mathbb{L}^2_{N+1}$ in the following way: The mapping $\mathcal{Q}_{N+1} : L^2\big((\mathbb{R}^3 \times \{\pm\frac{1}{2}\})^{N+1}\big) \to \mathbb{L}^2_{N+1}$ from (2.8) is used to define that for $\Psi \in \mathbb{L}^2_N$,

$$a_f^\dagger \Psi \;:=\; (N!)^{-\frac{1}{2}} \mathcal{Q}_{N+1}\big(f \otimes \Psi\big). \tag{2.16}$$

The adjoint of the creation operator $a_f^\dagger : \mathbb{F} \to \mathbb{F}$ of $f$ is termed the annihilation operator or annihilator $a_f : \mathbb{F} \to \mathbb{F}$ of $f$.

The creation operator $a_f^\dagger$ is a bounded operator on $\mathbb{F}$ (with $\|a_f^\dagger\| = \|f\|$). In particular, the adjoint of the annihilator $a_f$ is indeed $a_f^\dagger$ as indicated by the notation. For annihilators and creators, there hold the important *anticommutator relations* [21]: denoting $[A, B]_+ = AB + BA$, one has

$$[a_f, a_g]_+ = 0, \qquad [a_f^\dagger, a_g^\dagger]_+ = 0, \tag{2.17}$$

and if $f, g \in L^2(\mathbb{R}^3 \times \{\pm\frac{1}{2}\})$ are orthogonal, then

$$[a_f, a_g^\dagger]_+ = [a_f^\dagger, a_g]_+ = 0. \tag{2.18}$$

Furthermore, all creation and annihilation operators are nilpotent,

$$a_f a_f = a_f^\dagger a_f^\dagger = 0. \tag{2.19}$$

For the construction of the CC method, one restricts oneself to creation and annihilation operators belonging to the basis functions from the one-particle basis $B$ fixed above: For $f = \chi_P \in B$, let us denote $a_P := a_{\chi_P}$, $a_P^\dagger := a_{\chi_P}^\dagger$ for sake of brevity.

Using this notation, the reference determinant can then be written as

$$\Psi_0 \;=\; \Psi[1, \ldots, N] \;=\; \bigwedge_{I \in \mathrm{occ}} \chi_I \;=\; a_1^\dagger a_2^\dagger \ldots a_N^\dagger |\rangle; \tag{2.20}$$

also, we obtain from the anticommutator relations (2.17), (2.18) and the nilpotency (2.19) that for $P \in \mathbb{N}$ and arbitrary $k$-particle Slater determinants $\Psi_\mu = \Psi[P_1, \ldots, P_k]$

$$a_P \Psi_\mu = \begin{cases} 0 & \text{if } \; P \notin \mu \\ (-1)^i \Psi[P_1, .., P_{i-1}, P_{i+1}, .., P_k] & \text{if } P = P_i. \end{cases} \tag{2.21}$$

(Note that $\Psi[P_1, .., P_{I-1}, P_{I+1}, .., P_k]$ is a Slater determinant of $k-1$ functions – $a_P$ has "annihilated" the function $\chi_P$.) For $a_P^\dagger$, there holds

$$a_P^\dagger \Psi_\mu = \begin{cases} 0 & \text{if } P \in \mu \\ \Psi[P, P_1, \ldots, P_k] & \text{if } P \notin \mu \end{cases} \tag{2.22}$$

so that $a_P^\dagger$ "creates" the orbital $\chi_P$ in $\Psi[P_1, \ldots, P_k]$, making it a $(k+1)$-particle Slater determinant (unless $P$ is already an orbital in $\Psi[P_1, \ldots, P_k]$).

**Excitation operators and Slater bases.** The annihilation and creation operators are in particular the building blocks of excitation operators that will themselves be used to define the cluster operators used in CC theory. Their definition is as follows.

**Definition 2.4** (excitation operators)**.** For $k \leq N$ and any selection

$$I_1 < \ldots < I_k \in \mathrm{occ}, \qquad A_1 < \ldots < A_k \in \mathrm{virt}$$

of indices, we define the corresponding excitation operator by

$$X^{A_1, \ldots, A_k}_{I_1, \ldots, I_k} = a_{A_1}^\dagger a_{I_1} a_{A_2}^\dagger a_{I_2} \ldots a_{A_k}^\dagger a_{I_k}. \tag{2.23}$$

To characterize the action of excitation operators, we start by investigating how $X_I^A = a_A^\dagger a_I$ for $I \in \text{occ}, A \in \text{virt}$ acts on the reference determinant $\Psi_0 = \Psi[1,\dots,N]$: from the relations (2.21), (2.22), one finds that

$$X_I^A \Psi[1,\dots,I-1,I,I+1\dots,N] \quad = \quad \Psi[1,\dots,I-1,A,I+1,\dots,N],$$

*i.e.* $X_I^A$ replaces the occupied one-particle function $\chi_I$ in $\Psi_0$ by the virtual function $\chi_A$. More globally, an excitation operator $X_{I_1,\dots,I_k}^{A_1,\dots,A_k}$ with $I_1 < \dots < I_k \in \text{occ}$, $A_1 < \dots < A_k \in \text{virt}$ maps the reference to

$$X_{I_1,\dots,I_k}^{A_1,\dots,A_k}\Psi_0 \quad = \quad \Psi_\mu = \Psi[P_1,\dots,P_N] \quad \text{with} \quad P_I = \begin{cases} I & if \quad I \notin \{I_1,\dots,I_k\} \\ A_j & if \ I = I_j \in \{I_1,\dots,I_k\} \end{cases} \tag{2.24}$$

that is, to the Slater determinant with multi-index $\mu = [P_1,\dots,P_N]$ obtained from the reference index $[1,\dots,N]$ by replacing the occupied orbitals $I_1,\dots,I_k$ by the virtual orbitals $A_1,\dots,A_k$ (putting $A_i$ in the position of $I_i$). Let us collect all these multi-indices $\mu$, obtained by replacing $k$ occupied orbitals ($1 \le k \le N$) in the reference index $[1,\dots,N]$ by $k$ virtual indices in ascending order, in an index set $\mathcal{M}$. The set

$$\mathbb{B} = \{\Psi_\mu \mid \mu \in \mathcal{M}\} := \{\Psi_\mu \mid \Psi_\mu \ = \ X_{I_1,\dots,I_k}^{A_1,\dots,A_k}\Psi_0 \ \text{ for some exc. operator } \ X_{I_1,\dots,I_k}^{A_1,\dots,A_k}\}$$

is then a (so-called Slater) basis for the space $\mathbb{H}^{1,\perp} := \mathbb{H}^1 \cap \{\Psi_0\}^\perp$, *cf. e.g.* Lemma 1.17 in [36]. We thus have a one-to-one correspondence between the excitation operators $X_{I_1,\dots,I_k}^{A_1,\dots,A_k}$ and the elements of the basis $\mathbb{B}$. In the following analysis, it will be useful not to use the notation for excitation operators introduced above, but to index them by the multi-index $\mu$ defined by

$$X_\mu := X_{I_1,\dots,I_k}^{A_1,\dots,A_k} \qquad \text{if} \qquad X_{I_1,\dots,I_k}^{A_1,\dots,A_k}\Psi_0 = \Psi_\mu, \qquad \mu \in \mathcal{M}$$

and we will exclusively use the notation $X_\mu, \mu \in \mathcal{M}$ in the following, so that

$$\mathbb{B} = \{\Psi_\mu \mid \mu \in \mathcal{M}\} = \{X_\mu\Psi_0 \mid \mu \in \mathcal{M}\}. \tag{2.25}$$

To avoid confusion, we give the correspondence between the notation $X_\mu, \mu \in \mathcal{M}$ and the notation $X_{I_1,\dots,I_k}^{A_1,\dots,A_k}$ explicitly: for given $\mu \in \mathcal{M}$, the indices of the annihilation operators contained in the representation (2.23) are given by the set $\text{occ}\backslash\text{occ}(\mu)$ (indices not contained in $\mu$ must have been annihilated), while the indices of creation operators in (2.23) are the virtual orbitals contained in $\mu$, *i.e.* given by $\text{virt}(\mu)$. *Vice versa*, or $X_{I_1,\dots,I_k}^{A_1,\dots,A_k}$ given, the corresponding index $\mu = [P_1,\dots,P_N]$ can be determined by replacing in $[1,\dots,N]$ the indices $I_1,\dots,I_k$ by $A_1,\dots,A_k$ as above (or by using (2.24)).

**Some properties of excitation operators.** It is an essential fact for Coupled Cluster theory that excitation operators commute as a consequence of the relations (2.17), (2.18):

$$X_\alpha X_\beta = X_\beta X_\alpha. \tag{2.26}$$

The result of the above product is either also an excitation operator (up to a sign), or zero; more precisely, we have the following statements that follow directly from the definition (2.23) and the nilpotency property (2.19).

**Lemma 2.5.** *There holds $X_\alpha X_\beta \neq 0$ if and only if*

$$\text{virt}(\alpha) \cap \text{virt}(\beta) = \emptyset, \qquad \text{occ}(\alpha) \cup \text{occ}(\beta) = \text{occ}; \tag{2.27}$$

*in this case $X_\alpha X_\beta = \pm X_\mu$ with*

$$\text{virt}(\mu) = \text{virt}(\alpha) \ \dot\cup \ \text{virt}(\beta), \quad \text{occ}(\mu) = \text{occ}(\alpha) \cap \text{occ}(\beta) \tag{2.28}$$

*(where $\dot\cup$ denotes the disjoint union). In particular, for fixed $\alpha, \mu \in \mathcal{M}$, there holds*

$$X_\alpha X_\beta = \pm X_\mu \ \text{ for some } \ \beta \in \mathcal{M} \quad \text{ only if } \quad \text{virt}(\alpha) \subseteq \text{virt}(\mu), \text{occ}(\alpha) \supseteq \text{occ}(\mu); \tag{2.29}$$

*in this case, $\beta$ is unique and given by*

$$\text{virt}(\beta) = \text{virt}(\mu)\backslash\text{virt}(\alpha), \quad \text{occ}(\beta) = \text{occ}(\mu) \cup \big(\text{occ}\backslash\text{occ}(\alpha)\big). \tag{2.30}$$

Note also that the effect of application of $X_\alpha$ to a basis function $\Psi_\beta = X_\beta \Psi_0$, $\alpha, \beta \in \mathcal{M}$, is by the above rules either zero, or, if (2.27) is fulfilled, equals $\pm\Psi_\mu$ with $\mu$ determined by (2.28). Similar rules holds for the formal adjoints of excitation operator. If $X_\mu$ is given by (2.23), its adjoint is the so-called de-excitation operator

$$X_\mu^\dagger \;=\; a_{I_1}^\dagger a_{A_1} a_{I_2}^\dagger a_{A_2} \dots a_{I_r}^\dagger \dots a_{A_r};$$

for instance, there holds $X_\mu^\dagger \; X_\mu \; \Psi_0 \;=\; \Psi_0$. For later purposes, we fix the following more global analogue of Lemma 2.5.

**Lemma 2.6.** *For $\alpha, \beta \in \mathcal{M}$, we have $X_\alpha^\dagger X_\beta \Psi_0 \neq 0$ only if*

$$\mathrm{virt}(\alpha) \subseteq \mathrm{virt}(\beta), \qquad \mathrm{occ}(\alpha) \supseteq \mathrm{occ}(\beta).$$

*In particular, $\beta, \mu \in \mathcal{M}$ fixed, there holds $X_\alpha^\dagger X_\beta \Psi_0 = \pm X_\mu \Psi_0$ for some $\alpha \in \mathcal{M}$ only if*

$$\mathrm{virt}(\mu) \subseteq \mathrm{virt}(\beta), \qquad \mathrm{occ}(\mu) \supseteq \mathrm{occ}(\beta). \tag{2.31}$$

*In this case, $\alpha$ is unique and given by*

$$\mathrm{virt}(\alpha) = \mathrm{virt}(\beta) \backslash \mathrm{virt}(\mu), \quad \mathrm{occ}(\alpha) = \mathrm{occ}(\beta) \cup \big(\mathrm{occ} \backslash \mathrm{occ}(\mu)\big). \tag{2.32}$$

The statements of Lemma 2.6 follow easily by noting that for $\alpha, \beta, \mu \in \mathcal{M}$, Lemma 2.5 implies that $X_\alpha^\dagger X_\beta \; \Psi_0 = \pm X_\mu \; \Psi_0$ holds if and only if $X_\beta \; \Psi_0 = \pm X_\alpha X_\mu \; \Psi_0$. The assertion thus follows directly from (2.27)–(2.30).

**Cluster operators and their $\mathbb{L}^2$-continuity.** Using the representation (2.25) for the tensor basis $\mathbb{B}$ of $\mathbb{H}^{1,\perp}$, every intermediately normed function $\Psi = \Psi_0 + \Psi^* \in \mathbb{L}^2$ can be expanded as

$$\Psi \;=\; \Psi_0 + \Psi^* \;=\; \Psi_0 + \sum_{\alpha \in \mathcal{M}} s_\alpha X_\alpha \Psi_0 \;=:\; (I + S_{\Psi^*})\Psi_0 \tag{2.33}$$

of at most $N$-fold excitations $X_\alpha \Psi_0$ of the reference determinant $\Psi_0 \in \mathbb{B}$, where the operator $S_{\Psi^*}$ is a linear operator, so far well-defined on $\mathrm{span}\{\Psi_0\}$. We now approach a first step to the continuous CC formulation by proving that $S_{\Psi^*}$ uniquely defines a bounded operator on $\mathbb{L}^2$, the so-called *cluster operator* defined by $\Psi^* \in \mathbb{L}^2$. We also show that the definition of this operator (which we also denote by $S_{\Psi^*}$) is independent of the chosen one-particle basis $B$, but only depends on the reference subspace $R$, and add some statements on the $\mathbb{L}^2$-adjoint $S^\dagger$ of $S$ needed later. In contrast to the technical problems that arise in establishing the crucial properties of $S_{\Psi^*}$ as operator on $\mathbb{H}^1$ in the next sections, the proof is relatively harmless and follows mostly from arguments for the discrete, projected CC method, [40], Lemma 4.13.

**Theorem 2.7** ($\mathbb{L}^2$-continuity of the cluster operator; independence of the basis)**.** *Let $\Psi^* \in \{\Psi_0\}^\perp$. Then there holds the following:*

(*i*) *The cluster operator $S = S_{\Psi^*}$ defined* via (2.33) *possesses a unique continuous extension to $\mathbb{L}^2$,*

$$S \in B(\mathbb{L}^2, \mathbb{L}^2),$$

*the* cluster operator *defined by $\Psi^*$. There is a constant $a > 0$ such that*

$$\|\Psi^*\| \;\leq\; \|S\|_{\mathbb{L}^2 \to \mathbb{L}^2} \;\leq a\; \|\Psi^*\|. \tag{2.34}$$

(*ii*) *This operator $S \in B(\mathbb{L}^2, \mathbb{L}^2)$ is independent of the choice of the bases $B_{\mathrm{occ}}, B_{\mathrm{virt}}$ for $R, R^\perp$ and thus of the basis $\mathbb{B}$ for $\{\Psi_0\}^\perp$ in the following sense:*
*Let $\tilde{B} := \{\tilde{\chi}_I \mid I \in \mathrm{occ}\} \cup \{\tilde{\chi}_A \mid A \in \mathrm{virt}\}$ any one-particle basis for which*

$$\mathrm{span}\{\tilde{\chi}_I | I \in \mathrm{occ}\} = R, \qquad \mathrm{span}\{\tilde{\chi}_A | A \in \mathrm{virt}\} = R^\perp,$$

*and let* $\tilde{X}_\alpha, \alpha \in \mathcal{M}$ *according excitation operators and* $\tilde{\Psi}_\alpha = \tilde{X}_\alpha \Psi_0, \alpha \in \mathcal{M}$ *according basis functions of a basis* $\tilde{\mathbb{B}}$ *for* $\{\Psi_0\}^\perp$. *Then if*

$$\Psi^* \;=\; \sum_{\alpha\in\mathcal{M}} s_\alpha \Psi_\alpha \;=\; \sum_{\alpha\in\mathcal{M}} \tilde{s}_\alpha \tilde{\Psi}_\alpha,$$

*there holds*

$$S \;=\; \sum_{\alpha\in\mathcal{M}} s_\alpha X_\alpha \;=\; \sum_{\alpha\in\mathcal{M}} \tilde{s}_\alpha \tilde{X}_\alpha.$$

(*iii*) *The* $\mathbb{L}^2$*-adjoint* $S^\dagger$ *of* $S$ *is given by*

$$S^\dagger = \sum_{\mu\in\mathcal{M}} s_\mu X_\mu^\dagger \;\in\; B(\mathbb{L}^2, \mathbb{L}^2) \tag{2.35}$$

*and can also be bounded by*

$$\|\Psi^*\| \;\leq\; \|S^\dagger\|_{\mathbb{L}^2\to\mathbb{L}^2} \;\leq a\;\; \|\Psi^*\|. \tag{2.36}$$

*Proof.* In [40], Lemma 4.13, it was proved that for all linear combinations

$$\Psi^* = \sum_{\alpha\in\mathcal{M}} s_\alpha \Psi_\alpha, \quad \Psi = \sum_{\beta\in\mathcal{M}} c_\beta \Psi_\alpha \in\; \mathbb{L}^2, \tag{2.37}$$

of elements from $\mathbb{B}$ with only finitely many coefficients $s_\alpha, c_\beta$ nonzero, there holds

$$\|S\Psi\| \leq C_N \|\Psi^*\| \|\Psi\| \tag{2.38}$$

where $C_N$ is a constant independent of the number of terms in the expansion (2.37) of $\Psi^*, \Psi$. For all finite linear combinations $\Psi^*$, one can therefore in a first step use (2.38) to uniquely extend $S = S_{\Psi^*}$ to a continuous operator on $\mathbb{L}^2$ by taking $\mathbb{L}^2$-limits with respect to $\Psi$. In a second step, we write arbitrary $\Psi^* \in \mathbb{L}^2$ as limit of a sequence $\Psi_n$ of finite linear combinations. $S = S_{\Psi^*}$ can for such $\Psi^*$ be defined as the unique limit of the $B(\mathbb{L}^2, \mathbb{L}^2)$-Cauchy sequence $(S_{\Psi_n})_{n\in\mathbb{N}}$. In particular, the bound (2.38) also holds for $\Psi^*$, implying the upper bound in (2.34). Observing $\|S\Psi_0\| = \|\Psi^*\|$ provides the lower bound on $\|S\|_{\mathbb{L}^2\to\mathbb{L}^2}$. The rather straightforward proof of (ii) rewrites the annihilators and creators belonging to $\tilde{B}$ as sums of those built from $B$, see [36] for the details. Item (iii) follows by writing $S$ as a limit of finite sums of excitation operators and using the continuity of the adjoint mapping $\mathcal{A} : \mathbb{L}^2 \to \mathbb{L}^2, A \mapsto A^\dagger$. □

**Remark on the choice of one-particle bases.** Let us emphasize here that (ii) implies in particular that *any* bases for the occupied and virtual spaces $R, R^\perp$ might be used as long as the $L^2$-orthogonality between $B_{\text{occ}}$ and $B_{\text{virt}}$ is maintained. Therefore, the present analysis covers not only the canonical choice, where $B_{\text{occ}}$ is given by the first $N$ eigenfunctions of some discrete or continuous Fock operator, complemented by an orthogonal basis of $R^\perp$, but also many of the more sophisticated CC schemes which are not directly based on canonical orbitals (*i.e.* eigenfunctions of the Fock operator) anymore, but use certain localization criteria to rotate the occupied orbitals (to *e.g.* Foster–Boys-type orbitals [8], Pipek–Mazay-type orbitals [33] or enveloped localized orbitals [1]), or use non-orthogonal bases for the complement $B_{\text{virt}}$ (*e.g.* the projected atomic orbitals (PAOs) in the LCCSD approach [20, 41]). Also, the analysis can be extended to methods enhancing the virtual space by specialized basis functions taking the numerically hazardous electron-electron cusp [19, 24] into account (as *e.g.* the recent powerful $r_{1,2}$- and $f_{1,2}$- methods [25]). Note though that for non-orthogonal orbitals, the simple relations (2.21), (2.22) do not hold anymore, and some special care is required when implementing the action of annihilation and creation operators.

**Complete CI in terms of cluster operators and the problem with continuous CC.** We can use $S$ to rephrase the Schrödinger equation in the following way: $\Psi = \Psi_0 + \sum_{\alpha\in M} s_\alpha \Psi_\alpha$ solves the electronic Schrödinger equation in its intermediate normalisation formulation iff the corresponding cluster operator $S = S_{\Psi^*}$ solves

$$\langle \Phi, H(I+S)\Psi_0\rangle \quad = \quad E\,\langle \Phi, (I+S)\Psi_0\rangle \qquad \text{for all} \;\; \mu \in \mathcal{M}. \tag{2.39}$$

In finite dimensional spaces, (2.25) supplies a one-to-one correspondence between all functions $\Psi_0 + \Psi^* = \Psi_0 + \sum_{\alpha\in\mathcal{M}} s_\alpha \Psi_\alpha$ and the set of all operators $I + S = I + \sum_{\alpha\in\mathcal{M}} s_\alpha X_\alpha$, and the latter can then be expressed equivalently as the set of matrix exponentials

$$\{e^T \mid T \text{ is a cluster operator}\}, \tag{2.40}$$

giving rise to the reformulation of (2.39) in terms of CC theory as outlined in the introduction. In infinite-dimensional spaces, we have just proved that any $\Psi^* \in \mathbb{L}^2 \cap \{\Psi_0\}^\perp$ corresponds uniquely to a cluster operator $S \in B(\mathbb{L}^2, \mathbb{L}^2)$, and using standard Banach algebra theory, we may rewrite the set of operators $I + S$ (with $\mathbb{L}^2$-continuous $S$) by a set (2.40) of exponentials as in the finite dimensional case.

The problems stem from the plain fact that not all $\mathbb{L}_2$-convergent series $\Psi^* = \sum_{\mu\in\mathcal{M}} \Psi_\mu$ of basis functions are possible solutions because of the restriction that $\underline{\Psi} = \Psi_0 + \Psi^* \in \mathbb{H}^1$. We will therefore show, and that is in fact the main problem in establishing the continuous CC method, that $\Psi^* \in \{\Psi_0\}^\perp$ is in $\mathbb{H}^{1,\perp}$ if and only if its cluster operator is in $B(\mathbb{H}^1, \mathbb{H}^1)$. The set of all admissible cluster operators is then a closed subalgebra of $B(\mathbb{H}^1, \mathbb{H}^1)$, and we obtain the correspondence

$$\begin{aligned}\{\Psi_0 + \Psi^* \mid \Psi^* \in \mathbb{H}^{1,\perp}\} \quad &\simeq \quad \{I + S \mid S \in B(\mathbb{H}^1, \mathbb{H}^1) \;\text{ is a cluster operator}\} \\ &\simeq \{e^T \mid T \in B(\mathbb{H}^1, \mathbb{H}^1) \;\text{ is a cluster operator}\}\end{aligned}$$

that will allow to formulate the Coupled Cluster equations in analogy to their finite dimensional version in Section 5. To write out CC in its linked formulation, we will also have to prove the $\mathbb{H}^1$-continuity of the $\mathbb{L}^2$-adjoint $S^\dagger$ of a cluster operator $S \in B(\mathbb{H}^1, \mathbb{H}^1)$. The proof of these continuity statements is subject of the following two sections: the following section provides some technical preparations; then, the the main statements are proved in Section 4.

## 3. An equivalent norm on $\mathbb{H}^1$

As remarked in the last section, the crucial part in setting up the continuous CC method consists in proving that for $\Psi^* \in \mathbb{H}^{1,\perp}$, $S, S^\dagger$ defined in the last section are continuous as operators mapping $\mathbb{H}^1 \to \mathbb{H}^1$. Our proof of this fact relies on a specific representation of a norm $\|\cdot\|_{R^N}$ that is equivalent to the Sobolev norm on $\mathbb{H}^1$ induced by (2.3) on the one hand, but allows us to take advantage of the $L^2$-orthogonality between occupied and virtual orbitals on the other. The norm $\|\cdot\|_{R^N}$ will be introduced in this section, and the representation used in the next section will be the main result of this section, given in Proposition 3.4.

To begin with, let us note that the Sobolev inner product (2.3) on $\mathbb{H}^1$ may be viewed as being induced by the one-particle operator $1 + \Delta$ on $H^1$, viz., by constructing from it $1 + \Delta$ the many-particle operators $(1 + \Delta)_{x_i}$ in (2.2)[4]. In the same way, every one-particle operator $A$ inducing an inner product equivalent to that on $H^1$ can be used to define an inner product $\langle\langle \cdot, \cdot \rangle\rangle_N$ on $\mathbb{H}^1$ as that induced by the lifted operator $A := \sum_{i=1}^N (A)_{x_k}$, and it is not hard to check that then also $\langle\langle \cdot, \cdot \rangle\rangle_N$ and $\langle \cdot, \cdot \rangle_1$ are equivalent on $\mathbb{H}^1$, see *e.g.* [36], Lemma 1.25. This is used in the next Lemma to construct a useful norm $\|\cdot\|_{R^N}$ on $\mathbb{H}^1$.

**Definition/Lemma 3.1** ($R$-norm on $H^1$, $R^N$-norm on $\mathbb{H}^1$)**.** *Denote by $P$, $P^\perp$ the $L^2$-projectors on the subspace $R$ (Def. 2.1) and on its $L^2$-orthogonal complement $R^\perp$, respectively.*

(*i*) *On $H^1 \times H^1$, the operator*

$$P(1-\Delta)P + P^\perp(1-\Delta)P^\perp \tag{3.1}$$

[4]For antisymmetric $\Psi$, this is akin to the fact that $\|\Psi\|_1 = \mathrm{tr}(1-\Delta)\gamma^1$, where $\gamma^1$ the one-particle density matrix of $\Psi$ and the Laplacian acts on $x$. In fact arguments similar those used when dealing with these quantities will enter below in the computation of the projection on the $R^1$-orthonormal basis.

*defines an inner product*

$$\langle \varphi, \psi \rangle_R := \langle P\varphi, P\psi \rangle_1 + \langle P^\perp \varphi, P^\perp \psi \rangle_1 \tag{3.2}$$

*equivalent to the canonical $H^1$-inner product $\langle \cdot, \cdot \rangle_1$.*

(*ii*) *For any $\varphi \in R, \psi \in R^\perp \cap H^1$,*

$$\langle \varphi, \psi \rangle_R \;=\; 0. \tag{3.3}$$

(*iii*) *The quadratic form*

$$\langle \cdot, \cdot \rangle_{R^N} : \quad \mathbb{H}^1 \times \mathbb{H}^1 \to \mathbb{R}, \qquad \langle \Phi, \Psi \rangle_{R^N} \;=\; \textstyle\sum_{i=1}^N \langle \Phi, \Psi \rangle_{R_i}, \tag{3.4}$$

*associated with the operator*

$$\sum_{i=1}^N R_i \quad := \quad \sum_{i=1}^N (P(1-\Delta)P + P^\perp(1-\Delta)P^\perp)_{x_i} \tag{3.5}$$

*(constructed in the fashion outlined prior to this definition) is an inner product equivalent to the canonical induced inner product* (2.3) *on $\mathbb{H}^1$, and the induced norm $\|\cdot\|_{R^N}$ is equivalent to the canonical norm $\|\cdot\|_{\mathbb{H}^1}$ on $\mathbb{H}^1$.*

*Proof.* $P$ is of finite rank and therefore continuous as mapping $H^1 \to H^1$; we thus also have continuity of $P^\perp$ and therefore that

$$\|\varphi\|_R^2 := \|P\varphi\|_1^2 + \|P^\perp\varphi\|_1^2 \leq (\|P\|_{H^1 \to H^1}^2 + \|P^\perp\|_{H^1 \to H^1}^2)\|\varphi\|_1^2.$$

*Vice versa*, denoting the $L_2$-norm by $\|\cdot\|$,

$$\begin{aligned} \|\varphi\|_1^2 &\leq \|\varphi\|^2 + \|\nabla P\varphi\|^2 + \|\nabla P^\perp\varphi\|^2 + 2\|\nabla P\varphi\|\|\nabla P^\perp\varphi\| \\ &\leq \|\varphi\|^2 + 2(\|\nabla P\varphi\|^2 + \|\nabla P^\perp\varphi\|^2) \leq 2\|\varphi\|_R^2. \end{aligned}$$

holds; item (ii) is straightforward, and using the equivalence of the inner products from (i), (iii) readily follows with the remarks made ahead of this lemma. □

For antisymmetric functions $\Psi \in \mathbb{H}^1$, the expression (3.4) becomes a little simpler.

**Lemma 3.2.** *Let $\Psi \in \mathbb{H}^1$, then*

$$\|\Psi\|_{R^N} = N^{\frac{1}{2}} \|\Psi\|_{R_1}, \tag{3.6}$$

*with the norm $\|\Psi\|_{R_1}$ induced by the inner product $\langle \cdot, \cdot \rangle_{R_1} := \langle R_1 \cdot, \cdot \rangle$ associated with the operator $R_1$ from* (3.5) *(acting only on $x_1$).*

*Proof.* For fixed $i \in \{1, \ldots, N\}$, let us set as $\tilde{\Psi}$ the function which arises by exchanging the coordinates belonging to the index 1 and $i$. Then, for the components of the inner product induced by (3.5), $\langle R_i\Psi, \Psi \rangle = \langle R_1\tilde{\Psi}, \tilde{\Psi} \rangle = \langle \Psi, \Psi \rangle_{R_1}$, where the last equality follows from the antisymmetry of $\Psi$. The asserted statement follows. □

The above lemma puts us into the position of expressing the $R^N$-norm on $\mathbb{H}^1$ in terms of projection on a basis $\overline{\mathbb{B}}$ that is orthonormal with respect to the $R_1$-inner product. This basis is constructed in the next definition.

**Definition/Lemma 3.3.** *Let us choose an $\langle\cdot,\cdot\rangle_R$-orthonormal basis $\{\overline{\chi}_P \mid P \in \mathbb{N}\}$ of $H^1$ for which*

$$\mathrm{span}\{\overline{\chi}_I \mid I = 1,\dots,N\} \;=\; R, \qquad \overline{\mathrm{span}\{\overline{\chi}_A \mid A > N\}} \;=\; R^\perp.$$

*Also, denote by $\mathcal{N} \subseteq \mathbb{N}^{N-1}$ the set of those multi-indices $\overline{\nu} = (P_1,\dots,P_{N-1})$ of length $N-1$ that can be obtained from some $\mu \in \mathcal{M}$ by leaving out one of the indices (and keeping the order of the rest). The set*

$$\mathbb{B}_{N-1} \;=\; \{\Psi_{\overline{\nu}} := \mathcal{Q}\big(\otimes_{i=1}^{N-1} \chi_{P_k}\big) \mid \overline{\nu} = (P_1,\dots,P_{N-1}) \in \mathcal{N}\}$$

*is a basis of the $N-1$ electron space $\mathbb{L}^2_{N-1}$, and the set*

$$\overline{\mathbb{B}} \;:=\; \{\Phi_{(P,\overline{\nu})} := \overline{\chi}_P \otimes \Psi_{\overline{\nu}} \;\mid\; P \in \mathcal{J}, \overline{\nu} \in \mathcal{N}\}$$

*is an $R_1$-orthonormal system for which $\mathbb{H}^1 \subseteq \overline{\mathrm{span}\overline{\mathbb{B}}}$.*

*Proof.* We only prove the statements for $\overline{\mathbb{B}}$; the rest is straightforward. Obviously, $\overline{\mathbb{B}}$ is orthonormal in the $\langle\cdot,\cdot\rangle_{R^1}$-inner product. To prove $\mathbb{H}^1 \subseteq \overline{\mathrm{span}\overline{\mathbb{B}}}$, let a basis function $\Psi_\mu = \Psi[P_1,\dots,P_N] \in \mathbb{B}$ be given. By dividing the $N!$ summands of the Slater determinant $\Psi_\mu$ up into $N$ parts according to which of the functions $\chi_{P_i}$ turns up in the first position, we can write $\Psi_\mu$ as

$$\Psi_\mu \;=\; \frac{1}{(N!)^{\frac{1}{2}}} \sum_{\pi\in S(N)} (-1)^{|\pi|}\, \chi_{\pi(P_1)} \otimes \dots \otimes \chi_{\pi(P_N)} \;=\; N^{-\frac{1}{2}} \sum_{i=1}^{N} (-1)^i\, \chi_{P_i} \otimes \Psi_{\overline{\mu}_i}, \tag{3.7}$$

where $\overline{\mu}_i \in \mathcal{N}$ contains all indices from $\mu$ except the index $P_i$, so that $\Psi_{\overline{\mu}_i}$ is an $(N-1)$-electron Slater determinant from $\mathbb{B}_{N-1}$. We thus have $\Psi_\mu \in \mathrm{span}\overline{\mathbb{B}}$ and the assertion follows. □

Using the previous construction, the following lemma provides the following important working expression for the norm $\|\Psi\|_{R^N}$ of a wave function $\Psi$ which will later allow for the desired estimate for cluster operators. This expression may look somewhat complicated (incorporating antisymmetry and Slater-Condon rules), but will turn out useful later.

**Proposition 3.4** ($R^N$-norm of antisymmetric functions)**.** *For any function $\Psi \in \mathbb{H}^1$, $\Psi = d_{[1,\dots,N]}\Psi_0 + \sum_{\mu\in\mathcal{M}} d_\mu \Psi_\mu$, there holds*

$$\begin{aligned}\|\Psi\|^2_{R^N} = N^{\frac{1}{2}} \sum_{\Phi_{(P,\overline{\nu})}\in\overline{\mathbb{B}}} \langle \Phi_{(P,\overline{\nu})}, \Psi\rangle^2_{R^1} &= \sum_{I\in\mathrm{occ}} \sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{J\in\mathrm{occ}} (-1)^{|J,\overline{\nu}|-1}\; d_{[J,\overline{\nu}]}\; \gamma_{I,J} \right|^2 \\ &+ \sum_{A\in\mathrm{virt}} \sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{B\in\mathrm{virt}} (-1)^{|B,\overline{\nu}|-1} d_{[B,\overline{\nu}]} \gamma_{A,B} \right|^2 .\end{aligned} \tag{3.8}$$

*where, for sake of brevity, we introduced some notation:*

- *For $P \in \mathbb{N}$, $\overline{\nu} \in \mathcal{N}$, we denote by $[P,\overline{\nu}] \in \mathcal{M}$ the multi-index of length $N$ that contains all the $N-1$ indices from $\overline{\nu}$ and the index $P$ (in the order enforced by the requirement that $[P,\overline{\nu}] \in \mathcal{M}$).*
- *If $P$ is one of the indices in $\overline{\nu}$, we set $d_{[P,\overline{\nu}]} = 0$ for convenience.*
- *$|P,\overline{\nu}|$ denotes the position of the index $P$ in the multi-index $[P,\overline{\nu}] \in \mathcal{M}$.*
- *$\gamma_{P,Q} := \langle \overline{\chi}_P, \chi_Q\rangle_R$ for $P,Q \in \mathbb{N}$.*

*Proof.* Using Lemmas 3.2, 3.3, the first equality is clear by expansion of $\Psi$ in the $R_1$-orthonormal system $\overline{\mathbb{B}}$. For the second, we rewrite the middle term as

$$\|\Psi\|_{R^N}^2 = N^{\frac{1}{2}} \sum_{P\in\mathbb{N}} \sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{\mu\in\mathcal{M}} d_\mu \langle \Phi_{(P,\overline{\nu})}, \Psi_\mu \rangle_{R_1} \right|^2 \tag{3.9}$$

and compute for fixed $(P,\overline{\nu}) \in \mathbb{N}\times\mathcal{N}$ and $\mu = [P_1,\dots,P_N] \in \mathcal{M}$ the coefficient $\langle \Phi_{(\overline{\nu},P)}, \Psi_\mu\rangle_{R_1}$: By standard, Slater–Condon-like arguments [21], $\langle \Phi_{(P,\overline{\nu})}, \Psi_\mu\rangle_{R_1} \neq 0$ only if $\mu = [Q,\overline{\nu}]$ for some $Q\in\mathbb{N}$. In this case, the splitting (3.7) shows that

$$\langle \Phi_{(P,\overline{\nu})}, \Psi_\mu\rangle_{R_1} = N^{-\frac{1}{2}}(-1)^{|P,\overline{\nu}|} \langle \Phi_{(P,\overline{\nu})},\ \chi_Q \otimes \Phi_{\overline{\nu}}\rangle = N^{-\frac{1}{2}}(-1)^{|P,\overline{\nu}|-1} \langle \overline{\chi}_P, \chi_Q\rangle_R$$

because only one of the $N$ terms in (3.7) survives being tested with $\Phi_{(P,\overline{\nu})}$. Inserting this in (3.9) yields

$$\|\Psi\|_{R^N}^2 = \sum_{P\in\mathbb{N}} \sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{Q\in\mathbb{N}} (-1)^{|P,\overline{\nu}|-1} d_{[P,\overline{\nu}]} \gamma_{P,Q} \right|^2$$

which falls apart into two sums by the orthogonality condition (3.3), so one obtains the desired expression (3.8) □

## 4. Continuity properties of cluster operators

The present section approaches the main problem in the formulation of continuous CC, *i.e.*, the continuity properties of cluster operators belonging to functions $\Psi^* \in \mathbb{H}^{1,\perp} := \mathbb{H}^1 \cap \{\Psi_0\}^\perp$. The following theorem states the main result. Afterwards, we will at first comment on the difficulties that have to be approached and devise a roadmap to the proof given here. The actual proof of Theorem 4.1 is then covered in the rest of this section.

**Theorem 4.1** ($\mathbb{H}^1$-continuity of the operators $S, S^\dagger$)**.** *Let $R$ and thus $\Psi_0$ be fixed. By Theorem 2.7, every $\Psi^* \in \{\Psi_0\}^\perp \cap \mathbb{L}^2$ uniquely defines an according cluster operator $S = S_{\Psi^*} \in B(\mathbb{L}^2,\mathbb{L}^2)$.*

*(i) There holds*

$$\Psi^* \in \mathbb{H}^1 \quad\Leftrightarrow\quad S \in B(\mathbb{H}^1,\mathbb{H}^1), \tag{4.1}$$

*and for some constant $b_1 > 0$ independent of $\Psi^*$ that*

$$\|\Psi^*\|_{\mathbb{H}^1} \quad\leq\quad \|S\|_{\mathbb{H}^1\to\mathbb{H}^1} \quad\leq\quad b_1\|\Psi^*\|_1. \tag{4.2}$$

*(ii) For all $\Psi^* \in \{\Psi_0\}^\perp \cap \mathbb{L}^2$, there holds for the $\mathbb{L}^2$-adjoint $S^\dagger$ of $S$ that*

$$S^\dagger \in B(\mathbb{H}^1,\mathbb{H}^1), \qquad \|S^\dagger\|_{\mathbb{H}^1\to\mathbb{H}^1} \quad\leq\quad b_2\|\Psi^*\|. \tag{4.3}$$

*for some constant $b_2 > 0$.*

Note that the adjoint $S^\dagger$ is bounded in terms of the $\mathbb{L}_2$-norm of $\Psi^*$. In particular, the norm of $S^\dagger$ as mapping $\mathbb{H}^1 \to \mathbb{H}^1$ is not uniformly bounded from below by the $\mathbb{H}^1$-norm of $\Psi^*$ which is easily seen by choosing a sequence $\Psi_n^*$ for which $\|\Psi_n^*\|_1 = 1$ but $\|\Psi_n^*\| \to 0$; there then holds $\|T^\dagger_{\Psi_n^*}\|_1/\|\Psi_n^*\|_1 \leq \|\Psi_n^*\|/\|\Psi^*\|_1 \to 0$.

In contrast to the proof for $\mathbb{L}_2$-continuity (Thm. 2.7) and although the relation (2.33) between a function $\Psi$ and it s cluster operator is bluntly simple, the relations (4.1) to (4.3) for $S$ and $S^\dagger$ are rather hard to verify: the operator $S$ is easily seen to be non-compact in general, and to the authors knowledge, there are no investigations of the analytical properties of cluster operators available in the literature, except for those for finite-dimensional, "projected" CC analysed by Schneider in [40]. A direct transfer of the approaches taken there fails due to various

technical obstacles arising in the continuous case: To show $\mathbb{H}^1$-continuity, a complete eigenbasis of the discrete Fock operator of the system was used in [40], which does not have to exist anymore in the continuous case; also, the discretized Hamiltonian boundedly maps to $\ell_2$ for each Galerkin discretization, so that for analysis of the *discrete* Coupled Cluster equations, the need to show the continuity of $S^\dagger$ as mapping $\mathbb{H}^1 \to \mathbb{H}^1$ could be avoided. This is not the case anymore in the continuous setting. Note also that the continuity of $S : \mathbb{H}^1 \to \mathbb{H}^1$ only implies the continuity of its $\mathbb{H}^1$-adjoint $S^{\dagger,\mathbb{H}^1} : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$, but not the $\mathbb{H}^1$-continuity of the restriction of the $\mathbb{L}^2$-adjoint $S^\dagger : \mathbb{L}^2 \to \mathbb{L}^2$ to $\mathbb{H}^1$.

Nevertheless, our proof given here will utilize a major argument already used in [40]: One of the main observations made there is that due to the nilpotency of excitation operators, the number of indices $\alpha, \beta \in \mathcal{M}$ for which $X_\alpha X_\beta = X_\mu$ for some fixed $\mu \in \mathcal{M}$ is bounded by a constant $C_N$ independent of $\mu$, allowing for an estimate of the $\ell_1$-norms of parts of $\|S\Psi\|_{L_2}$ by the $\ell_2$-norm. We will use analogous arguments here; in fact, the main reason for using for the complicated norm expression for $\|\cdot\|_{R^N}$ derived in the last section is it that allows the application of such estimates in order to isolate the different terms of the norm representation of $\|\Psi\|_{R^N}, \|\Psi^*\|_{R^N}$ from the expression for $\|S\Psi\|_{R^N}$.

**Organization and overview of the proof.** The proof is organized in three parts: The first two sections prove the statements (i) on $S$ and (ii) on $S^\dagger$, respectively; to convey the readability of the proof, two lemmas will be stated and used during the proofs, but proved only afterwards in the separate third section. Before we start with (i), we sketch for the convenience of the reader how we show the hardest part, namely that $\Psi^* \in \mathbb{H}^{1,\perp}$ implies $S \in B(\mathbb{H}^1, \mathbb{H}^1)$: We found in Section 3 that the norm $\|\cdot\|_{R^N}$ introduced there is equivalent to the Sobolev norm on $\mathbb{H}^1$; therefore, $\Psi^* \in \mathbb{H}^{1,\perp}$ fixed, $S \in B(\mathbb{H}^1, \mathbb{H}^1)$ holds if we can show that for all $\Psi \in \{\Psi_0\}^\perp \cap \mathbb{H}^1$,

$$\|S\Psi\|_{R^N} \leq C\|\Psi^*\|_{R^N}\|\Psi\|_{R^N}, \tag{4.4}$$

with $C$ a constant independent of $\Psi^*$ (Note that $S$ is bounded by $\|\Psi^*\|_1$ on $\operatorname{span}\{\Psi_0\}$.) We recall the expression (3.8) for the $\|\cdot\|_{R^N}$-norm computed in Proposition 3.4,

$$\|S\Psi\|_{R^N}^2 = \underbrace{\sum_{\substack{I\in\mathrm{occ}\\ \overline{\nu}\in\mathcal{N}}} \left| \sum_{J\in\mathrm{occ}} (-1)^{|J,\overline{\nu}|-1}\, d_{[J,\overline{\nu}]}\, \gamma_{I,J} \right|^2}_{\text{(I)}} + \underbrace{\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}} \left| \sum_{B\in\mathrm{virt}} (-1)^{|B,\overline{\nu}|-1} d_{[B,\overline{\nu}]}\gamma_{A,B} \right|^2}_{\text{(II)}}, \tag{4.5}$$

in which the norm is expressed in terms of sums over indices $\mu = [P, \overline{\nu}] \in \mathcal{M}$, *i.e.* $\mu \in \mathcal{M}$ consists of the occupied or virtual index $P$ and of those contained in $\overline{\nu}$, an ordered index belonging to an $(N-1)$-particle Slater determinant. (Remember also the convention $d_{[P,\overline{\nu}]} = 0$ if $P$ is contained in the indices of $\overline{\nu}$.) The first term (I) allows for a simple estimate based on Theorem 2.7. To outline the proceeding for the term (II), we denote

$$\Psi^* = \sum_{\alpha\in\mathcal{M}} s_\alpha \Psi_\alpha, \quad \Psi = \sum_{\beta\in\mathcal{M}} c_\beta \Psi_\beta, \quad S\Psi = \sum_{\mu\in\mathcal{M}} d_\mu X_\mu \Psi_0 = \sum_{\alpha\in\mathcal{M}}\sum_{\beta\in\mathcal{M}} s_\alpha c_\beta X_\alpha X_\beta \Psi_0. \tag{4.6}$$

and then use two auxiliary statements: *(L1)* bases on Lemma 2.5 to give an expression for coefficients $d_\mu$ *via* determining the "fitting" combinations of $\alpha, \beta$; the resulting terms tailored for the norm representation (4.5) will look a bit more complicated than in Lemma 2.5, but are determined by the same rules. Counting the number of combinations $\alpha, \beta$ that amount to a fixed $\mu$ then gives rise to an estimate *(L2)*, the modification of the estimate from [40] mentioned above, and the estimate (4.4) can finally be proved. As noted above, the proofs of *(L1)*, *(L2)* will be given in the third part of the proof. The proof of (ii) for the adjoint of $S$ will follow similar lines as that of (i), and analogous statements *(L1')*, *(L2')* will enter the proof and also be verified only afterwards.

*Proof of (i).* **Continuity and estimates for $S$.**

We start by noting that if $S\Psi_0 = \Psi^* \notin \mathbb{H}^{1,\perp}$, then $S \notin B(\mathbb{H}^1, \mathbb{H}^1)$; it therefore remains to prove the converse. Let $\Psi^* \in \mathbb{H}^{1,\perp}$; following the above outline, we start to show (4.4) by estimating (I) on the right hand side of (4.5): Application of the Cauchy–Schwarz inequality to the sum over $J \in \mathrm{occ}$ yields

$$\sum_{\substack{I\in\mathrm{occ}\\ \overline{\nu}\in\mathcal{N}}} \left| \sum_{J\in\mathrm{occ}} (-1)^{|J,\overline{\nu}|-1}\, d_{[J,\overline{\nu}]}\, \gamma_{I,J} \right|^2 \leq \left( \sum_{I,J\in\mathrm{occ}} \gamma_{I,J}^2 \right) \left( \sum_{J\in\mathrm{occ}} \sum_{\overline{\nu}\in\mathcal{N}} d_{[J,\overline{\nu}]}^2 \right). \tag{4.7}$$

Noting

$$\sum_{I,J\in\mathrm{occ}} \gamma_{I,J}^2 = \sum_{I\in\mathrm{occ}} \|\chi_I\|_R^2 = \|\Psi_0\|_{R^N}, \qquad \sum_{J\in\mathrm{occ}} \sum_{\overline{\nu}\in\mathcal{N}} d_{[J,\overline{\nu}]}^2 \;\leq\; N \cdot \sum_{\mu\in\mathcal{M}} d_\mu^2 \;=\; N \cdot \|S\Psi\|$$

shows with the $\mathbb{L}^2$-continuity estimate (2.34) for $S$ that

$$\text{(I)} \leq aN \|\Psi_0\|_{R^N} \|\Psi^*\| \|\Psi\|.$$

To obtain (4.4), we now show that

$$\text{(II)} \;=\; \sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}} \left| \sum_{B\in\mathrm{virt}} (-1)^{|B,\overline{\nu}|-1} d_{[B,\overline{\nu}]} \gamma_{A,B} \right|^2$$

in (4.5) is also bounded by the $\|\cdot\|_{R^N}$-norms of $\Psi$ and $\Psi^*$. To do so, we express the coefficients $d_\mu = d_{[B,\overline{\nu}]}$ in terms of the coefficients $t_\alpha, c_\beta$. To denote this in a concise manner, we define the following.

**Definition 4.2.** We define a partial ordering on $\mathcal{M}$ by writing $\overline{\alpha} \prec \overline{\nu}$ iff both $\mathrm{occ}(\overline{\nu}) \subset \mathrm{occ}(\overline{\alpha})$ and $\mathrm{virt}(\overline{\alpha}) \subset \mathrm{virt}(\overline{\nu})$ holds. Two indices $\overline{\alpha} \prec \overline{\nu}$ given, this uniquely defines a corresponding multi-index $\overline{\nu} \ominus \overline{\alpha} \in \mathcal{M}$ of length $N$ by

$$\mathrm{occ}(\overline{\nu} \ominus \overline{\alpha}) = (\mathrm{occ} \backslash \mathrm{occ}(\overline{\alpha})) \,\dot\cup\, \mathrm{occ}(\overline{\nu}), \qquad \mathrm{virt}(\overline{\nu} \ominus \overline{\alpha}) = \mathrm{virt}(\overline{\nu}) \backslash \mathrm{virt}(\overline{\alpha}).$$

With that, we obtain the announced representation for the coefficients $d_{[B,\overline{\nu}]}$ in terms of the coefficients $c_\alpha, t_\alpha$ of $\Psi, \Psi^*$, see Lemma 4.3 in the "lemma section" below for the proof.

***(L1)*** *For*

$$S\Psi = \sum_{\mu\in\mathcal{M}} d_\mu X_\mu \Psi_0 = \sum_{\alpha\in\mathcal{M}} \sum_{\beta\in\mathcal{M}} s_\alpha c_\beta X_\alpha X_\beta \Psi_0$$

*from (4.6), the coefficients $d_{[B,\overline{\nu}]}$, $B \in \mathrm{virt}, \overline{\nu} \in \mathcal{N}$, are given by*

$$d_{[B,\overline{\nu}]} = \sum_{\substack{\alpha\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}} (-1)^{[B,\overline{\nu}]-1} (-1)^{[B,\overline{\alpha}]-1} \sigma_{(\overline{\nu},\overline{\alpha})} \big( s_{[B,\overline{\alpha}]}\; c_{\overline{\nu}\ominus\overline{\alpha}} \;+\; c_{[B,\overline{\alpha}]} s_{\overline{\nu}\ominus\overline{\alpha}} \big), \tag{4.8}$$

*where $\sigma_{(\overline{\nu},\overline{\alpha})}$ is a sign distribution only depending on $(\overline{\nu}, \overline{\alpha})$.*

Inserting this into (II), we obtain in terms of $t_\alpha, c_\beta$ that

$$\text{(II)} = \sum_{A\in\mathrm{virt}} \sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{B\in\mathrm{virt}} \sum_{\substack{\alpha\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}} (-1)^{|B,\overline{\alpha}|-1} \sigma_{(\overline{\nu},\overline{\alpha})} \left( t_{[B,\overline{\alpha}]}\; c_{\overline{\nu}\ominus\overline{\alpha}} \;+\; c_{[B,\overline{\alpha}]} t_{\overline{\nu}\ominus\overline{\alpha}} \right) \gamma_{A,B} \right|^2.$$

We split this up into

$$\text{(II)} \;\le\; 2\left( \sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{M}}} \left| \sum_{\substack{\alpha\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})} \left( \sum_{B\in\text{virt}} (-1)^{|B,\overline{\alpha}|-1}\; t_{[B,\overline{\alpha}]}\; \gamma_{A,B} \right) c_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \right. \tag{4.9}$$

$$\left. + \sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{M}}} \left| \sum_{\substack{\alpha\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})} \left( \sum_{B\in\text{virt}} (-1)^{|B,\overline{\alpha}|-1}\; c_{[B,\overline{\alpha}]}\gamma_{A,B} \right) t_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \right). \tag{4.10}$$

To estimate the two terms in (4.9) resp. (4.10), we now use the following estimate (4.11), which we will prove in Lemma 4.4 below.

***(L2)*** *For any sequences* $(b_{[A,\overline{\nu}]}) \in \ell_2(\mathcal{N}\times\text{virt}), (c_\beta)\in\ell_2(\mathcal{M})$ *and for any sign distribution* $\sigma_{(\overline{\nu},\overline{\alpha})}$ *there holds that*

$$\sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{N}}} \left| \sum_{\substack{\overline{\alpha}\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})}\; b_{[A,\overline{\alpha}]}\; c_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \;\le\; C_N \;\; \|(b_{[A,\overline{\nu}]})\|_2^2 \;\; \|(c_\beta)\|_2^2 \tag{4.11}$$

*with the constant* $C_N$ *only depending on the fixed number* $N$ *of electrons.*

We apply *(L2)* to $(c_\beta)$ and to

$$(b_{[A,\overline{\alpha}]}) \;:=\; \left( \sum_{B\in\text{virt}} (-1)^{|B,\overline{\alpha}|-1}\; t_{[B,\overline{\alpha}]}\; \gamma_{A,B} \right)$$

and obtain

$$\sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{M}}} \left| \sum_{\overline{\alpha}\prec\overline{\nu}} \sigma_{(\overline{\nu},\overline{\alpha})} \left( \sum_{B\in\text{virt}} (-1)^{|B,\overline{\alpha}|-1}\; t_{[B,\overline{\alpha}]}\; \gamma_{A,B} \right) c_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \le C_N \left( \sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{M}}} \left| \sum_{B\in\text{virt}} (-1)^{|B,\overline{\nu}|-1}\; t_{[B,\overline{\alpha}]}\gamma_{A,B} \right|^2 \right) \|(c_\beta)\|_2^2$$

$$\le C_N \;\; \|\Psi^*\|_{R^N} \|\Psi\|$$

by comparison of the bracketed term in the middle with the expression for the $R^N$-norm of $\Psi^*$. Proceeding with the summand in line (4.10) in the same way gives the other way around

$$\sum_{\substack{A\in\text{virt}\\ \overline{\nu}\in\mathcal{M}}} \left| \sum_{\overline{\alpha}\prec\overline{\nu}} \sigma_{(\overline{\nu},\overline{\alpha})} \left( \sum_{B\in\text{virt}} (-1)^{|B,\overline{\alpha}|-1}\; c_{[A,\overline{\alpha}]}\gamma_{A,B} \right) t_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \;\le\; C_N \;\; \|\Psi^*\| \cdot \|\Psi\|_{R^N}.$$

Thus altogether, $\|S\Psi\|_{R^N}$ is bounded up to a constant by $\|\Psi^*\|_{R^N}\|\Psi\|_{R^N}$ ,

$$\|S\Psi\|_{R^N} \lesssim \|\Psi^*\|_{R^N} \cdot \|\Psi\|_{R^N},$$

which implies $S\in B(\mathbb{H}^1,\mathbb{H}^1)$ and the upper estimate for $\|S\|_{\mathbb{H}^1\to\mathbb{H}^1}$ in (4.2). Observing $\|S\|_{\mathbb{H}^1\to\mathbb{H}^1} \ge \|S\Psi_0\|_1 = \|\Psi^*\|_1$ finishes the proof of the statements on $S$.

*Proof of (ii).* **Continuity of $S^\dagger$**

The proof of the $\mathbb{H}^1$-continuity of $S^\dagger$ is similar to that for $S$, we therefore only sketch the proceeding. Denoting

$$S^\dagger\Psi \;=\; \sum_{\mu\in\mathcal{M}} e_\mu\Psi_\mu \;=\; \sum_{\alpha\in\mathcal{M}}\sum_{\beta\in\mathcal{M}} s_\alpha c_\beta X_\alpha^\dagger X_\beta\Psi_0,$$

one again uses the representation (3.8) to estimate $\|S^\dagger\Psi\|_{R^N}$; one obtains

$$\|S^\dagger\Psi\|_{R^N}^2 = \underbrace{\sum_{\substack{I\in\mathrm{occ}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{J\in\mathrm{occ}}(-1)^{|J,\overline{\nu}|-1}\; e_{[J,\overline{\nu}]}\;\gamma_{I,J}\right|^2}_{\text{(I')}} + \underbrace{\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{B\in\mathrm{virt}}(-1)^{|B,\overline{\nu}|-1}e_{[B,\overline{\nu}]}\gamma_{A,B}\right|^2}_{\text{(II')}}. \tag{4.12}$$

The term (I') can be bounded in terms of the $\mathbb{L}^2$-norms of $\Psi,\Psi^*$ in the same fashion as (I) in the proof for $S$, only that now the $\mathbb{L}^2$-estimate (2.36) for $S^\dagger$ is used:

$$\text{(I)} \;\le\; N\|\Psi_0\|\;\|S^\dagger\Psi\| \;\le\; aN\|\Psi_0\|_{R^N}\;\|\Psi\|\|\Psi^*\|.$$

Next, (II') can be estimated in terms of $\|\Psi\|_{R^N}$ and $\|\Psi^*\|$. We again express the coefficients $e_{[B,\overline{\nu}]}$ *via* the "fitting" combinations of $s_\alpha, c_\beta$. For the adjoint, there holds the following statement (see Lem. 4.3):

***(L1')*** *For $B\in\mathrm{virt},\overline{\nu}\in\mathcal{N}$,*

$$e_{[B,\overline{\nu}]} \;=\; \sum_{\substack{\overline{\beta}\in\mathcal{N}\\ \overline{\nu}\prec\overline{\beta}}}(-1)^{|B,\overline{\beta}|-1}(-1)^{|B,\overline{\nu}|-1}\sigma_{(\overline{\beta},\overline{\nu})}s_{\overline{\beta}\ominus\overline{\nu}}c_{[\overline{\beta},B]} \tag{4.13}$$

*with some sign distribution $\sigma_{(\overline{\beta},\overline{\nu})}$ depending only on $(\overline{\beta},\overline{\nu})$.* □

Inserting this into (II') gives

$$\text{(II')} = \sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{\substack{\overline{\beta}\in\mathcal{N}\\ \overline{\nu}\prec\overline{\beta}}}\sigma_{(\overline{\beta},\overline{\nu})}\left(\sum_{B\in\mathrm{virt}}(-1)^{|B,\overline{\beta}|-1}c_{[\overline{\beta},B]}\gamma_{A,B}\right)s_{\overline{\beta}\ominus\overline{\nu}}\right|^2$$

To estimate further, we use an analogue to the statement *(L2)* above (also see Lem. 4.4 below for the proof):

***(L2')*** *For any sequences $(b_{[A,\overline{\nu}]})\in\ell_2(\mathcal{N}\times\mathrm{virt}),(c_\beta)\in\ell_2(\mathcal{M})$ and any sign distribution $\sigma_{(\overline{\beta},\overline{\nu})}$,*

$$\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{\substack{\overline{\beta}\in\mathcal{N}\\ \overline{\nu}\prec\overline{\beta}}}\sigma_{(\overline{\beta},\overline{\nu})}\;b_{[A,\overline{\beta}]}\;c_{\overline{\beta}\ominus\overline{\nu}}\right|^2 \;\le\; C_N\;\;\|(b_{[A,\overline{\nu}]})\|_2^2\;\;\|(c_\beta)\|_2^2. \tag{4.14}$$

This yields

$$\text{(II')} \le C_N\left(\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{B\in\mathrm{virt}}(-1)^{|B,\overline{\nu}|-1}c_{[\overline{\nu},B]}\gamma_{A,B}\right|^2\right)\|(s_\beta)\|_2^2 = C_N\|\Psi\|_{R^N}^2\|\Psi^*\|^2$$

which implies the upper bound (4.3) for the $\mathbb{H}^1$-norm of $T^\dagger$. □

*Proof of the statements* (L) *used in the proof.* To tie up the proof of Theorem 4.1, we will now redeliver the proof of the statements *(L1), (L1'), (L2)* and *(L2')* used in the above. The following combinatorial lemma shows *(L1), (L1')*.

**Lemma 4.3.** *Let $\overline{\nu} \in \mathcal{N}$ be an ordered index belonging to an $(N-1)$-particle Slater determinant and $B \in$ virt not be contained in the indices of $\overline{\nu}$.*
*(i) For $\alpha, \beta \in \mathcal{M}$, there holds*

$$X_\alpha X_\beta = \pm X_{[B,\overline{\nu}]} \quad \Longleftrightarrow \quad \begin{cases} & \textit{Either} \\ \alpha & = [B,\overline{\alpha}] \textit{ with } \overline{\alpha} \prec \overline{\nu};\ \beta = \overline{\nu} \ominus \overline{\alpha} \\ & \textit{or} \\ \beta & = [B,\overline{\beta}] \textit{ with } \overline{\beta} \prec \overline{\nu};\ \alpha = \overline{\nu} \ominus \overline{\beta} \end{cases}$$

*We therefore have for*

$$S\Psi = \sum_{\mu \in \mathcal{M}} d_\mu X_\mu \Psi_0 = \sum_{\alpha \in \mathcal{M}} \sum_{\beta \in \mathcal{M}} s_\alpha c_\beta X_\alpha X_\beta \Psi_0$$

*from (4.6) that*

$$d_{[B,\overline{\nu}]} = \sum_{\overline{\alpha} \prec \overline{\nu}} \sigma^{\text{virt}}_{(\overline{\nu},\overline{\alpha},B)} \left( s_{[B,\overline{\alpha}]}\ c_{\overline{\nu} \ominus \overline{\alpha}} \quad + \quad c_{[B,\overline{\alpha}]} s_{\overline{\nu} \ominus \overline{\alpha}} \right) \tag{4.15}$$

*where $\sigma^{\text{virt}}$ is some sign distribution, dependent on $\overline{\nu}, \overline{\alpha}, B$.*
*(ii) The sign distribution $\sigma^{\text{virt}}$ from (i) can be written as*

$$\sigma^{\text{virt}}_{(\overline{\nu},\overline{\alpha},B)} = (-1)^{[B,\overline{\nu}]-1} (-1)^{[B,\overline{\alpha}]-1} \sigma_{(\overline{\nu},\overline{\alpha})}$$

*with $\sigma_{(\overline{\nu},\overline{\alpha})}$ a sign distribution only depending on $(\overline{\nu},\overline{\alpha})$.*
*(iii) For*

$$S^\dagger \Psi \ = \ \sum_{\mu \in \mathcal{M}} e_\mu \Psi_\mu \ = \ \sum_{\alpha \in \mathcal{M}} \sum_{\beta \in \mathcal{M}} s_\alpha c_\beta X_\alpha^\dagger X_\beta \Psi_0,$$

*there holds*

$$e_{[B,\overline{\nu}]} \ = \ \sum_{\substack{\overline{\beta} \in \mathcal{N} \\ \overline{\nu} \prec \overline{\beta}}} (-1)^{|B,\overline{\beta}|-1} (-1)^{|B,\overline{\nu}|-1} \sigma_{(\overline{\beta},\overline{\nu})} s_{\overline{\beta} \ominus \overline{\nu}} c_{[\overline{\beta},B]} \tag{4.16}$$

*with some sign distribution $\sigma_{(\overline{\beta},\overline{\nu})}$ depending only on $(\overline{\beta},\overline{\nu})$.*

*Proof.* Noting that either $\alpha$ or $\beta$ must contain the creator for $\chi_B$, the statements in (i) are a straightforward "translation" of (2.29), (2.30) to the notation of Definition 4.2. To show (ii), we determine the sign in $X_\alpha X_\beta = \pm X_{[B,\overline{\nu}]}$. Because excitation operators commute, we can assume without loss of generality that $\alpha$ contains the index $B$, $\alpha = [B,\overline{\alpha}]$ with $\overline{\alpha} \prec \overline{\nu}$. For the same reason, we can permute the excitation operators in $X_\alpha, X_\beta$ to obtain $X_\alpha = a_B^\dagger a_{J_1} X_{\hat{\alpha}}$, $X_{[B,\overline{\nu}]} = a_B^\dagger a_{J_2} X_{\hat{\mu}}$, with some $J_1, J_2 \in$ occ (the indices that "got excited" to $B$, respectively) and $X_{\hat{\alpha}}$, $X_{\hat{\mu}}$ some excitation operator or the identity. We note that

$$(-1)^{[B,\overline{\alpha}]-1} = (-1)^{J_1-1}, \quad (-1)^{[B,\overline{\nu}]-1} = (-1)^{J_2-1} \tag{4.17}$$

and compute

$$\sigma_{(\overline{\nu},\overline{\alpha},A)} = \langle X_\alpha X_\beta \Psi_0, X_{[B,\overline{\mu}]} \Psi_0 \rangle \ = \ \langle a_{J_1} X_{\hat{\alpha}} X_\beta \Psi_0, a_{J_2} X_{\hat{\mu}} \Psi_0 \rangle \ = \ (*).$$

The indices of $\beta$ are determined by (i) as $\overline{\nu} \ominus \overline{\alpha}$, those of $\hat{\alpha}$ are

$$\text{occ}(\hat{\alpha}) = \text{occ}(\overline{\alpha}) \ \dot\cup \ (\text{occ} \backslash \text{occ}(\beta)), \qquad \text{virt}(\hat{\alpha}) = \text{virt}(\overline{\alpha}).$$

We note that this uniquely determines $\hat{\alpha}$ and $\beta$ from $\overline{\alpha}$ and $\overline{\nu}$, establishing an injective mapping $(\overline{\alpha},\overline{\mu}) \mapsto (\hat{\alpha},\beta)$. We may thus permute indices such that $X_{\hat{\alpha}}X_\beta\Psi_0 = \sigma_{(\overline{\nu},\overline{\alpha})}\Psi_\nu$, where $\Psi_\nu \in \mathbb{B}$ is some ordered Slater determinant, and $\sigma_{(\overline{\nu},\overline{\alpha})}$ is a function of $\overline{\nu},\overline{\alpha}$. We thus obtain that $(*) = \sigma_{(\overline{\nu},\overline{\alpha})}\langle a_{J_1}\Psi_\nu, a_{J_2}\Psi_{\hat{\mu}}\rangle$. From (2.21),

$$a_{J_1}\Psi_\nu = (-1)^{J_1-1}\Psi_{\overline{\nu}}, \quad a_{J_2}X_{\hat{\mu}}\Psi_0 = (-1)^{J_2-1}\Psi_{\overline{\mu}}$$

with $\overline{\nu},\overline{\mu}$ the indices obtained by removing $J_1$ from $\nu$ resp. $J_2$ from $\hat{\mu}$. $\Psi_{\overline{\nu}},\Psi_{\overline{\mu}}$ are therefore Slater determinants from the ordered $(N-1)$-particle basis set $\mathbb{B}_{N-1}$ for $L^2_{N-1}$ introduced in Lemma 3.3. We continue the calculation by

$$\sigma_{(\overline{\nu},\overline{\alpha})}\langle a_{J_1}\Psi_\nu, a_{J_2}\Psi_{\hat{\mu}}\rangle = (-1)^{J_1-1}(-1)^{J_2-1}\sigma_{(\overline{\nu},\overline{\alpha})}\langle\Psi_{\overline{\nu}},\Psi_{\overline{\mu}}\rangle \quad = \quad (-1)^{J_1-1}(-1)^{J_2-1}\sigma_{(\overline{\nu},\overline{\alpha})};$$

the last step following because $\langle\Phi_{\overline{\nu}},\Phi_{\overline{\mu}}\rangle = \delta_{\overline{\nu},\overline{\mu}}$ for $\Psi_{\overline{\nu}},\Psi_{\overline{\mu}} \in \mathbb{B}_{N-1}$ and because the whole term must give $\pm 1$ by (i). Using (4.17), this completes the proof of (ii). To prove (iii), one uses (2.31), (2.32) to determine the coefficients $\alpha,\beta$ for which $X^\dagger_\alpha X_\beta\Psi_0 = X_{[B,\overline{\nu}]}\Psi_0$ by Lemma 2.6 in analogy to the proceeding in (i). One finds that in this case, for some $\overline{\beta} \prec \overline{\nu}$, $\beta = [B,\overline{\beta}]$ and $\alpha = \overline{\beta} \ominus \overline{\nu}$ have to hold. It remains to compute the corresponding sign factor

$$\sigma_{(\overline{\beta},\overline{\mu},B)} = \langle X^\dagger_\alpha X_{[B,\overline{\beta}]}\Psi_0, X_{[B,\overline{\nu}]}\Psi_0\rangle \quad = \quad \langle X_\alpha X_{[B,\overline{\nu}]}\Psi_0, X_{[B,\overline{\beta}]}\Psi_0\rangle;$$

applying the argument given in (ii) to the right hand side shows that this can be written as $\sigma_{(\overline{\beta},\overline{\mu},B)} = \sigma_{(\overline{\beta},\overline{\mu})}(-1)^{|B,\overline{\beta}|-1}(-1)^{|B,\overline{\mu}|-1}$, completing the proof. □

The last ingredient that is missing for Theorem 4.1 is the proof of the estimates *(L2)*, *(L2')*, given in the next lemma. An analogue to the first estimate (4.18) was already proven in [40], where it was central to the analysis for the projected CC equations in the discrete setting. We prove our re-formulation for the present context with an improved constant and derive from it the estimate (4.19), which was used as *(L2)* to show continuity of $S^\dagger$.

**Lemma 4.4.** *For any sequences $(b_{[A,\overline{\nu}]}) \in \ell_2(\mathcal{N}\times\mathrm{virt}), (c_\beta) \in \ell_2(\mathcal{M})$, there holds for any sign distribution $\sigma_{(\overline{\nu},\overline{\alpha})}$ that*

$$\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{\substack{\overline{\alpha}\in\mathcal{N}\\ \overline{\alpha}\prec\overline{\nu}}}\sigma_{(\overline{\nu},\overline{\alpha})}\; b_{[A,\overline{\alpha}]}\; c_{\overline{\nu}\ominus\overline{\alpha}}\right|^2 \quad \le \quad C_N \quad \|(b_{[A,\overline{\nu}]})\|_2^2 \quad \|(c_\beta)\|_2^2 \tag{4.18}$$

*and also for any sign distribution $\sigma_{(\overline{\beta},\overline{\nu})}$ that*

$$\sum_{\substack{A\in\mathrm{virt}\\ \overline{\nu}\in\mathcal{N}}}\left|\sum_{\substack{\overline{\beta}\in\mathcal{N}\\ \overline{\nu}\prec\overline{\beta}}}\sigma_{(\overline{\beta},\overline{\nu})}\; b_{[A,\overline{\beta}]}\; c_{\overline{\beta}\ominus\overline{\nu}}\right|^2 \quad \le \quad C_N \quad \|(b_{[A,\overline{\nu}]})\|_2^2 \quad \|(c_\beta)\|_2^2. \tag{4.19}$$

*Proof.* We start by estimating for a fixed index $\overline{\nu} \in \mathcal{N}$ the number of indices $\overline{\alpha} \in \mathcal{N}$ for which $\overline{\alpha} \prec \overline{\nu}$ holds: By definition, $\overline{\alpha} \prec \overline{\nu}$ iff $\mathrm{virt}(\overline{\alpha}) \subseteq \mathrm{virt}\ (\overline{\nu})$ and $\mathrm{occ}(\overline{\nu}) \subseteq \mathrm{occ}\ (\overline{\alpha})$. Denoting by $r$ the number of virtuals contained in $\overline{\nu}$, the number of possible indices $\overline{\alpha} \prec \overline{\nu}$ containing $s$ virtuals ($s < r$ by definition) is given by $\binom{r}{s}\binom{N-((N-1)-r)}{r-s} = \binom{r}{s}\binom{r-1}{r-s}$. Summing up over all $s < r$ gives

$$\sum_{s=1}^{r-1}\binom{r}{s}\binom{r-1}{r-s} \quad \le \quad \sum_{s=1}^{N-2}\binom{N-1}{s}\binom{N}{N-s} \quad \le \quad \binom{2N}{N} \quad =: \quad C_N$$

using *e.g.* Vandermonde's identity. Now, we can estimate the left hand of (4.18) by noting that for every fixed $\overline{\nu}$, the sum over $\overline{\alpha} \prec \overline{\nu}$ contains at most $C_N$ non-null summands; thus for each $A \in \mathrm{virt}$,

$$\sum_{\overline{\nu}\in\mathcal{N}} \left| \sum_{\substack{\overline{\alpha}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})}\ b_{[A,\overline{\alpha}]}\ c_{\overline{\nu}\ominus\overline{\alpha}} \right|^2 \leq C_N \sum_{\overline{\nu}\in\mathcal{N}} \sum_{\substack{\overline{\alpha}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} b^2_{[A,\overline{\alpha}]} c^2_{\overline{\nu}\ominus\overline{\alpha}}$$

$$= C_N \sum_{\overline{\alpha}\in\mathcal{N}} \left( b^2_{[A,\overline{\alpha}]} \cdot \sum_{\substack{\overline{\nu}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} c^2_{\overline{\nu}\ominus\overline{\alpha}} \right) \leq C_N \ \Big( \sum_{\overline{\nu}\in\mathcal{N}} b^2_{[A,\overline{\nu}]} \Big)\ \|(c_\beta)_{\beta\in\mathcal{M}}\|^2_{\ell_2(\mathcal{M})},$$

In the last step, we used that for $\overline{\alpha} \neq \overline{\beta}$, $\overline{\nu} \ominus \overline{\alpha} \neq \overline{\nu} \ominus \overline{\beta}$ holds, which implies

$$\sum_{\substack{\overline{\nu}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} c^2_{\overline{\nu}\ominus\overline{\alpha}} \leq \sum_{\alpha\in\mathcal{M}} c^2_\alpha.$$

This shows (4.18). To prove (4.19), we note that (4.18) means that for fixed $(c_\alpha) \in \ell_2(\mathcal{M})$ and a given sign distribution $\sigma_{(\overline{\nu},\overline{\alpha})}$, the mapping

$$M: \quad \ell_2(\mathrm{virt}\times\mathcal{N}) \to \ell_2(\mathrm{virt}\times\mathcal{N}), \quad (e_{[A,\overline{\nu}]}) \ \mapsto \ \left( \sum_{\substack{\overline{\alpha}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})}\ e_{[A,\overline{\alpha}]}\ c_{\overline{\nu}\ominus\overline{\alpha}} \right)$$

is continuous with norm $\|M\| \leq C_N^{\frac{1}{2}} \|(c_\alpha)\|_2$. We compute the adjoint of $M$: Denoting the $\ell_2$-inner product on $\ell_2(\mathrm{virt}\times\mathcal{N})$ by

$$\langle (a_{[A,\overline{\nu}]}), (b_{[A,\overline{\nu}]}) \rangle \ = \ \langle (a_{[A,\overline{\nu}]}), (b_{[A,\overline{\nu}]}) \rangle_{\ell_2(\mathrm{virt}\times\mathcal{N})}$$

there holds for $(e_{[A,\overline{\nu}]}) \in \ell_2(\mathrm{virt}\times\mathcal{N})$ that

$$\langle M(e_{[A,\overline{\nu}]}), (b_{[A,\overline{\nu}]}) \rangle = \sum_{\substack{\overline{\nu}\in\mathcal{N} \\ A\in\mathrm{virt}}} \left( \sum_{\substack{\overline{\alpha}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})}\ e_{[A,\overline{\alpha}]}\ c_{\overline{\nu}\ominus\overline{\alpha}} \right) b_{[A,\overline{\nu}]}$$

$$= \sum_{\substack{\overline{\alpha}\in\mathcal{N} \\ A\in\mathrm{virt}}} \left( \sum_{\substack{\overline{\nu}\in\mathcal{N} \\ \overline{\alpha}\prec\overline{\nu}}} \sigma_{(\overline{\nu},\overline{\alpha})}\ b_{[A,\overline{\nu}]}\ c_{\overline{\nu}\ominus\overline{\alpha}} \right) e_{[A,\overline{\alpha}]} = \left\langle (e_{[A,\overline{\nu}]}), \left( \sum_{\substack{\overline{\beta}\in\mathcal{N} \\ \overline{\nu}\prec\overline{\beta}}} \sigma_{(\overline{\beta},\overline{\nu})}\ b_{[A,\overline{\beta}]}\ c_{\overline{\beta}\ominus\overline{\nu}} \right) \right\rangle,$$

where we re-labelled the indices $\overline{\nu} \to \overline{\beta}, \overline{\alpha} \to \overline{\nu}$ in the last step. $M^\dagger$ is therefore given by

$$M^\dagger: \ \ell_2(\mathrm{virt}\times\mathcal{N}) \to \ell_2(\mathrm{virt}\times\mathcal{N}), \quad (b_{[A,\overline{\nu}]}) \ \mapsto \ \left( \sum_{\substack{\overline{\beta}\in\mathcal{N} \\ \overline{\nu}\prec\overline{\beta}}} \sigma_{(\overline{\beta},\overline{\nu})}\ b_{[A,\overline{\beta}]}\ c_{\overline{\beta}\ominus\overline{\nu}} \right) \cdots$$

$M^\dagger$ is also continuous with $\|M^\dagger\| = \|M\| \leq C_N^{\frac{1}{2}} \|(c_\alpha)\|_{\ell_2}$, and writing this out gives (4.19). □

## 5. The continuous Coupled Cluster equations

As the main result of this work, we now set up the continuous version of the Coupled Cluster equations and show that they are equivalent the exact (weak) eigenproblem (1.1) in Theorem 5.3. To do so, we need two more lemmas. The first one uses Theorem 4.1 to extend cluster operators to $\mathbb{H}^{-1}$; the second is the continuous version of Lemma 4.2 and Theorem 4.3 in [40], giving the formal justification for the exponential parametrization used in CC theory.

**Lemma 5.1** (continuity of $S : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$)**.** *Let $\Psi^* = \sum_{\mu\in\mathcal{M}^*} s_\mu\Psi_\mu \in \{\Psi_0\}^\perp$. The corresponding cluster operator $S = S_{\Psi^*} \in B(\mathbb{L}^2,\mathbb{L}^2)$ can be extended to a continuous operator $S : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$. In particular, each excitation operator $X_\mu$ can be continuously extended to an operator $X_\mu : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$, and there holds $S = \sum_{\mu\in\mathcal{M}^*} s_\mu X_\mu$ in $B(\mathbb{H}^{-1},\mathbb{H}^{-1})$.*

*Proof.* By virtue of Theorem 4.1, the $\mathbb{L}^2$-adjoint $S^\dagger$ of $S$ is a bounded operator on $\mathbb{H}^1$. Its $\mathbb{H}^1$-adjoint $\tilde{S} : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$ is therefore also continuous with $\|\tilde{S}\|_{\mathbb{H}^{-1}\to\mathbb{H}^{-1}} = \|S^\dagger\|_{\mathbb{H}^1\to\mathbb{H}^1}$. We show that $\tilde{S}$ extends $S \in B(\mathbb{L}^2,\mathbb{L}^2)$ (and therefore may be denoted in a slight abuse of notation as $S$ as above): We identify $\Psi \in \mathbb{L}^2$ with the functional

$$F(\cdot) \;=\; \langle\Psi,\cdot\rangle \in (\mathbb{L}^2)' \;\subseteq\; \mathbb{H}^{-1}$$

and obtain the assertion from

$$\tilde{S}F \;:=\; F(S^\dagger\cdot) \;=\; \langle\Psi,S^\dagger\cdot\rangle \;=\; \langle S\Psi,\cdot\rangle.$$

Theorem 4.1 in particular implies that $X_\mu : \mathbb{H}^{-1} \to \mathbb{H}^{-1}$ is continuous and well-defined, and $S$ and $\sum_{\mu\in\mathcal{M}^*} s_\mu X_\mu$ coincide on the dense subset $\mathbb{L}^2$, so $S = \sum_{\mu\in\mathcal{M}} s_\mu X_\mu$ also follows. □

**Lemma 5.2** (properties of the exponential function on the algebra of cluster operators)**.**

*The set*

$$L := \left\{ t_0 I + S \mid t_0 \in \mathbb{R},\; S = \sum_{\alpha\in\mathcal{M}} s_\alpha X_\alpha \in B(\mathbb{H}^1,\mathbb{H}^1) \right\}$$

*is a closed commutative subalgebra of $B(\mathbb{H}^1,\mathbb{H}^1)$, of which each element with $t_0 \neq 0$ is invertible. The exponential function $\exp(X) = \sum_{i=0}^N X^i/i!$ is a local $C^\infty$-diffeomorphism mapping onto $L\backslash\{0\}$. In particular,* exp *is a bijection between the sets*

$$\mathcal{S} = \left\{ S \mid \; S = \sum_{\alpha\in\mathcal{M}} s_\alpha X_\alpha \in B(\mathbb{H}^1,\mathbb{H}^1) \right\} \quad \textit{and} \quad I + \mathcal{S} = \left\{ I + S \mid S = \sum_{\alpha\in\mathcal{M}} s_\alpha X_\alpha \in B(\mathbb{H}^1,\mathbb{H}^1) \right\}.$$

*The lemma also holds if $\mathbb{H}^1$ is replaced by $\mathbb{H}^{-1}$ in the above.*

*Proof.* Taking Theorem 4.1 into account, the proof for the properties of $L$ is identical with that from [40], Lemma 4.2, and Theorem 4.3. Because $L$ is a commutative subalgebra of $B(\mathbb{H}^1,\mathbb{H}^1)$ resp. $B(\mathbb{H}^{-1},\mathbb{H}^{-1})$, the exponential function is a local $C^\infty$-diffeomorphism on $L\backslash\{0\}$, see *e.g.* [39]. The series terminates at $i = N$ because any product of more than $N$ excitation operators contains more than $N$ annihilators for the $N$ occupied orbitals and thus has to vanish by the nilpotency (2.19) of annihilation operators. exp maps $\mathcal{S}$ to $I + \mathcal{S}$ by definition, and on $I + \mathcal{S}$, its inverse is given by the (terminating) logarithmic series $\log(X) = \sum_{i=1}^N (-1)^{i-1}(X - I)^i/i$ (see [40]), which obviously maps to $\mathcal{S}$, so the lemma is proved. □

**The continuous Coupled Cluster equations.** We are now in the position to show that the exact (weak) eigenproblem (1.1) is equivalent to the continuous Coupled Cluster equations formulated in the following theorem.

**Theorem 5.3.** *Let $\underline{\Psi} \in \mathbb{H}^1$ with $\langle\Psi_0,\underline{\Psi}\rangle \neq 0$ and $E \in \mathbb{R}$. The following are equivalent:*

*(i) $\underline{\Psi}$ solves the (weak, CI) eigenproblem*

$$\langle\Phi,\; (H - E)\underline{\Psi}\rangle \;=\; 0, \quad \textit{for all } \Phi \in \mathbb{H}^1. \tag{5.1}$$

*(ii) $\underline{\Psi} = e^T\Psi_0$ for some cluster operator $T \in B(\mathbb{H}^1,\mathbb{H}^1)$ which fulfils the (continuous) unlinked Coupled Cluster equations*

$$\langle\Phi,(H - E)e^T\Psi_0\rangle = 0, \quad \textit{for all } \; \Phi \in \mathbb{H}^1. \tag{5.2}$$

(*iii*) $\underline{\Psi} = e^T\Psi_0$ *for some cluster operator* $T \in B(\mathbb{H}^1, \mathbb{H}^1)$ *which fulfils the (continuous) linked Coupled Cluster equations,*

$$E \;=\; \langle \Psi_0, He^T\Psi_0\rangle, \qquad \langle \Psi_\mu, e^{-T}He^T\Psi_0\rangle = 0 \qquad \textit{for all} \;\; \mu \in \mathcal{M}. \tag{5.3}$$

The advantage of the linked formulation (5.3) over the unlinked CC equations (5.2) is that by using the properties of the algebra of annihilation and creation operators, the term $e^{-T}He^T$ can expanded into a *terminating* Baker–Campbell–Hausdorff series which can be evaluated exactly in a finite basis set. Therefore, (5.3) is the set of equations that is almost exclusively discretized and solved in practice.

Note in this context that the above equivalence of linked and unlinked formulation does not need to hold anymore if in a discretized setting, based on certain selection criteria, only some of the amplitudes of the discretised basis are used for a computation. In this case, $e^{T^\dagger}$ is not necessarily surjective anymore; to guarantee this, the set of selected amplitudes has to be *excitation complete*, which is for instance the case for canonical models like CCSD, CCSDT etc., see [40] for details.

*Proof.* Using Theorem 4.1, $\underline{\Psi} \in \mathbb{H}^1$ solves the set of equations (5.1) iff there is a continuous cluster operator $S \in B(\mathbb{H}^1, \mathbb{H}^1)$ such that $\underline{\Psi} = (I + S)\Psi_0$ and

$$\langle \Phi, (H - E)(I + S)\Psi_0\rangle = 0 \quad \text{for all} \;\; \Phi \in \mathbb{H}^1. \tag{5.4}$$

By Lemma 5.2, there is a unique cluster operator $T \in B(\mathbb{H}^1, \mathbb{H}^1)$ such that $I+S = e^T$, so that (5.4) is equivalent to finding $T \in B(\mathbb{H}^1, \mathbb{H}^1)$ such that

$$\langle \Phi, (H - E)e^T\Psi_0\rangle = 0, \quad \text{for all} \;\; \Phi \in \mathbb{H}^1, \tag{5.5}$$

or in other words, $0 = (H - E)e^T\Psi_0 \in \mathbb{H}^{-1}$. By Theorem 4.1, the $\mathbb{L}^2$-adjoint $T^\dagger$ of $T$ is continuous as mapping $\mathbb{H}^1 \to \mathbb{H}^1$; therefore, $e^{T^\dagger}$ is a continuous invertible mapping $\mathbb{H}^1 \to \mathbb{H}^1$, and (5.5) is equivalent to

$$\langle e^{-T^\dagger}\Phi, (H - E)e^T\Psi_0\rangle = 0, \quad \text{for all} \;\; \Phi \in \mathbb{H}^1.$$

Due to the continuity of the adjoint mapping $\mathcal{A} : A \mapsto A^\dagger$, we have

$$\langle e^{-T^\dagger}\Phi, (H - E)e^T\Psi_0\rangle \;=\; \langle \Phi, (e^{-T^\dagger})^\dagger(H - E)e^T\Psi_0\rangle \;=\; \langle \Phi, e^{-T}(H - E)e^T\Psi_0\rangle$$

with the exponential $e^{-T}$ of $-T$ taken in $\mathbb{H}^{-1}$, which is equivalent to (5.3). □

**The continuous Coupled Cluster function.** To treat the nonlinear linked CC equations (5.3) in the finite dimensional setting, these are in practice often comprised to a nonlinear function, the *Coupled Cluster function*

$$\mathbf{f} : \ell_2(\mathcal{M}_d) \to \ell_2(\mathcal{M}_d), \quad (t_\alpha) \mapsto \big(\langle \Psi_\alpha, e^{-T(t_\alpha)}He^{T(t_\alpha)}\Psi_0 \,\rangle\big)_{\alpha\in\mathcal{M}}, \tag{5.6}$$

with $T(t_\alpha) = \sum_{\alpha\in\mathcal{M}_d} t_\alpha X_\alpha$. Any coefficient vector $(t_\alpha)$ fulfilling $\mathbf{f}(t_\alpha) = 0$ then corresponds to a solution of the discrete Schrödinger equation *via* the relation $\Psi_d = e^{T(t_\alpha)}\Psi_0$. Normally, Newton-like methods are used to solve for $\mathbf{f}(t_\alpha) = 0$. We will now end this work by reformulating the continuous linked formulation (5.3) in terms of a nonlinear operator $f$, discretizations of which then yield the canonical discrete, projected Coupled Cluster equations (5.6), and this viewpoint will enable statements on the CC method in the framework of nonlinear functional analysis in a following publications.

In contrast to the finite dimensional case, we are faced with the restriction that the cluster operator $T$ sought in the CC equations (5.2), (5.3) has to be contained in $B(\mathbb{H}^1, \mathbb{H}^1)$ by Theorem 5.3. the consequence of this is that the domain of continuous Coupled Cluster function $f$ is not $\ell_2(\mathcal{M})$ (which would reflect the space $\mathbb{L}^2$), but only a closed subspace of it (reflecting the subspace $\mathbb{H}^1$). This space is introduced in the following.

**Definition 5.4** (The $H^1$-coefficient space $\mathbb{V}$)**.** On $\ell_2(\mathcal{M})$, we define

$$\left\langle (t_\alpha), (s_\alpha) \right\rangle_{\mathbb{V}} \;:=\; \Big\langle \sum_{\alpha\in\mathcal{M}} t_\alpha \Psi_\alpha, \sum_{\beta\in\mathcal{M}} s_\beta \Psi_\beta \Big\rangle_1^2, \qquad \|(t_\alpha)\|_{\mathbb{V}} \;:=\; \langle (t_\alpha), (t_\alpha) \rangle_{\mathbb{V}}^{1/2}, \tag{5.7}$$

and a subspace $\mathbb{V}$ of $\ell_2(\mathcal{M})$ by

$$\mathbb{V} \;:=\; \{ t \in \ell_2(\mathcal{M}) \mid \|t\|_{\mathbb{V}} < \infty \} \subseteq \ell_2(\mathcal{M}),$$

Denoting as above by $T(t_\alpha)$ the cluster operator defined by $(t_\alpha)$ and $\Psi(t_\alpha) := T(t_\alpha)\Psi_0$, there holds $(t_\alpha) \in \mathbb{V}$ if and only if $\Psi^*(t_\alpha) \in \mathbb{H}^{1,\perp}$. $(\mathbb{V}, \langle\cdot,\cdot\rangle_{\mathbb{V}})$ is thus complete and therefore a Hilbert space. Also, by Theorem 4.1, the linear mapping

$$\mathcal{T} \;:\; (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \to (\mathcal{S}, \|\cdot\|_{\mathbb{H}^1\to\mathbb{H}^1}), \qquad (t_\alpha) \;\mapsto\; T(t_\alpha) = \sum_{\alpha\in\mathcal{M}} t_\alpha X_\alpha,$$

is a bounded bijection (Here, $\mathcal{S} \subseteq B(\mathbb{H}^1, \mathbb{H}^1)$ is the algebra of $\mathbb{H}^1$-bounded cluster operators $S$ introduced in Lem. 5.2). In the same vein,

$$\mathcal{T}^\dagger \;:\; (\mathbb{V}, \|\cdot\|_{\mathbb{V}}) \to (\mathcal{S}^\dagger, \|\cdot\|_{\mathbb{H}^1\to\mathbb{H}^1}), \qquad (t_\alpha) \;\mapsto\; T^\dagger(t_\alpha) = \sum_{\alpha\in\mathcal{M}} t_\alpha X_\alpha^\dagger$$

is bounded. With this, we obtain the continuous version of the CC root equation $\mathbf{f}(t_\alpha) = 0$.

**Theorem 5.5** (the continuous CC function)**.** *A cluster operator $T = T(t_\alpha^*) \in B(\mathbb{H}^1, \mathbb{H}^1)$ fulfils the linked CC equations*

$$E \;=\; \langle \Psi_0, He^T\Psi_0 \rangle, \qquad \langle \Psi_\mu, e^{-T}He^T\Psi_0 \rangle = 0, \quad \textit{for all} \;\; \mu \in \mathcal{M} \tag{5.8}$$

*if and only if the corresponding coefficient vector $(t_\alpha^*) \in \mathbb{V}$ is a root of the continuous Coupled Cluster function*

$$f : \mathbb{V} \;\to\; \mathbb{V}', \quad (t_\alpha) \;\mapsto\; \big( \langle \Psi_\alpha, e^{-T(t_\alpha)}He^{T(t_\alpha)}\Psi_0 \rangle \big)_{\alpha\in\mathcal{M}}. \tag{5.9}$$

*that is, if it solves the nonlinear continuous CC operator equation*

$$f(t_\alpha^*) \;=\; 0 \;\in\; \mathbb{V}'. \tag{5.10}$$

*Proof.* The equivalence of (5.8) and (5.10) is straightforward with the remarks made beforehand. It remains to show that for fixed $(t_\alpha) \in \mathbb{V}$, $\langle f(t_\alpha), \cdot \rangle$ defines a continuous functional on $\mathbb{V}$ (denoting by $\langle\cdot,\cdot\rangle$ the usual $\ell_2(\mathcal{M})$-inner product). For $(s_\alpha) \in \mathbb{V}$, we obtain with the boundedness of the Hamiltonian [48], Theorem 4.1, Lemma 5.2 and equation (5.8) that

$$\begin{aligned} \langle f(t_\alpha), (s_\alpha) \rangle &= \sum_{\alpha\in\mathcal{M}} \langle s_\alpha \Psi_\alpha, e^{-T(t_\alpha)}He^{T(t_\alpha)}\Psi_0 \rangle \\ &\le \|T(s)\Psi_0\|_{\mathbb{H}^1} \|e^{-T(t_\alpha)}He^{T(t_\alpha)}\Psi_0\|_{\mathbb{H}^{-1}} \;\le\; C(t)\|s\|_{\mathbb{V}}, \end{aligned}$$

where the constant $C(t)$ depends on the $\mathbb{V}$-norm of $t$, proving the claim. □

## 6. Concluding remarks and outlook

By the virtue of Theorem 5.3, we have obtained the continuous Coupled Cluster equations (5.2), (5.3), which are (up to the very mild restrictions of intermediate normalization) equivalent to the original operator eigenvalue problem (1.1), the electronic Schrödinger equation. Moreover, exact eigenvectors of the eigenproblem for the Hamiltonian correspond to the solutions of the root equation (5.10) for the continuous CC function (5.9), which in the continuous context defines a nonlinear operator between the coefficient space $\mathbb{V}$ and its dual space $\mathbb{V}'$. The discrete CC equations (5.6) for a fixed basis set, normally used as starting point in quantum chemistry, can now be interpreted as a Galerkin discretization of the root equation for the CC function $f$, consisting in computation of an approximate solution $(t_\alpha)$ from a Galerkin subspace $\mathbb{V}_d \subseteq \mathbb{V}$ for which

$$\langle f(t_\alpha), (s_\alpha)\rangle = 0 \quad \text{for all} \quad (s_\alpha) \in \mathbb{V}_d. \tag{6.1}$$

In the forthcoming publication [38], we will use this approach to treat the CC function in the formalism of nonlinear operator analysis: based on a local strong monotonicity statement for the CC function $f$, we will be able to derive existence and uniqueness results for the solutions of the continuous and discrete CC equations, statements on the convergence of the discrete solutions towards the continuous solution and of discrete CC energies towards the true ground state energy $E$.

## References


[1] A.A. Auer and M. Nooijen, Dynamically screened local correlation method using enveloping localized orbitals. *J. Chem. Phys.* **125** (2006) 24104.

[2] R.J. Bartlett, Many-body perturbation theory and coupled cluster theory for electronic correlation in molecules. *Ann. Rev. Phys. Chem.* **32** (1981) 359.

[3] R.J. Bartlett and M. Musial, Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **79** (2007) 291.

[4] R.J. Bartlett and G.D. Purvis, Many-body perturbation theory, coupled-pair many-electron theory, and the importance of quadruple excitations for the correlation problem. *Int. J. Quantum Chem.* **14** (1978) 561.

[5] U. Benedikt, M. Espig, W. Hackbusch and A.A. Auer, Tensor decomposition in post-Hartree-Fock methods. I. Two-electron integrals and MP2. *J. Chem. Phys.* **134** (2011) 054118.

[6] F.A. Berezin, *The Method of Second Quantization*. Academic Press (1966).

[7] R.F. Bishop, An overview of coupled cluster theory and its applications in physics. *Theor. Chim. Acta* **80** (1991) 95.

[8] S.F. Boys, Construction of some molecular orbitals to be approximately invariant for changes from one molecule to another. *Rev. Mod. Phys.* **32** (1960) 296.

[9] A. Chamorro, Method for construction of operators in Fock space. *Pramana* **10** (1978) 83.

[10] O. Christiansen, Coupled cluster theory with emphasis on selected new developments. *Theor. Chem. Acc.* **116** (2006) 106.

[11] P.G. Ciarlet (Ed.) and C. Lebris (Guest Ed.), Handbook of Numerical Analysis X: Special Volume. *Comput. Chem.* Elsevier (2003).

[12] J. Čížek, Origins of coupled cluster technique for atoms and molecules. *Theor. Chim. Acta* **80** (1991) 91.

[13] F. Coerster, Bound states of a many-particle system. *Nucl. Phys.* **7** (1958) 421.

[14] F. Coerster and H. Kümmel, Short range correlations in nuclear wave functions. *Nucl. Phys.* **17** (1960) 477.

[15] *Computational Chemistry Comparison and Benchmark Data Base*. National Institute of Standards and Technology, available on `http://cccbdb.nist.gov/`

[16] T.D. Crawford and H.F. Schaeffer III, An introduction to coupled cluster theory for computational chemists. *Rev. Comput. Chem.* **14** (2000) 33.

[17] H.L. Cycon, R.G. Froese, W. Kirsch and B. Simon, Schrödinger Operators with Applications to Quantum Mechanics and Global Geometry, Series *Theor. Math. Phys.* Springer (1987).

[18] V. Fock, Konfigurationsraum und zweite Quantelung. *Z. Phys.* **75** (1932) 622.

[19] S. Fournais, M. Hoffmann-Ostenhof, T. Hoffmann-Ostenhof and T. Østergaard Sørensen, Sharp regularity results for Coulombic many-electron wave functions. *Commun. Math. Phys.* **255** (2005) 183.

[20] C. Hampel and H.-J. Werner, Local treatment of electron correlation in coupled cluster theory. *J. Chem. Phys.* **104** (1996) 6286.

[21] T. Helgaker, P. Jørgensen and J. Olsen, *Molecular Electronic-Structure Theory*. John Wiley & Sons (2000).

[22] P.D. Hislop and I.M. Sigal, Introduction to spectral theory with application to Schrödinger operators. *Appl. Math. Sci.* **113** Springer (1996).

[23] W. Hunziker and I.M. Sigal, The quantum N-body problem. *J. Math. Phys.* **41** (2000) 6.

[24] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics. *Commun. Pure Appl. Math.* **X** (1957) 151.

[25] W. Klopper, F.R. Manby, S. Ten no and E.F. Vallev, R12 methods in explicitly correlated molecular structure theory. *Int. Rev. Phys. Chem.* **25** (2006) 427.

[26] W. Kutzelnigg, Error analysis and improvement of coupled cluster theory. *Theor. Chim. Acta* **80** (1991) 349.

[27] W. Kutzelnigg, Unconventional aspects of Coupled Cluster theory, in *Recent Progress in Coupled Cluster Methods, Theory and Applications, Series: Challenges and Advances in Computational Chemistry and Physics* **11** (2010). To appear.

[28] H. Kümmel, Compound pair states in imperfect Fermi gases. *Nucl. Phys.* **22** (1961) 177.

[29] H. Kümmel, K.H. Lührmann and J.G. Zabolitzky, Many-fermion theory in expS- (or coupled cluster) form. *Phys. Rep.* **36** (1978) 1.

[30] T.J. Lee and G.E. Scuseria, Achieving chemical accuracy with Coupled Cluster methods, in *Quantum Mechanical Electronic Structure Calculations with Chemical Accuracy*, edited by S.R. Langhof. Kluwer Academic Publishers, Dordrecht (1995) 47.

[31] F. Neese, A. Hansen and D.G. Liakos Efficient and accurate approximations to the local coupled cluster singles doubles method using a truncated pair natural orbital basis. *J. Chem. Phys.* **131** (2009) 064103.

[32] M. Nooijen, K.R. Shamasundar and D. Mukherjee, Reflections on size-extensivity, size-consistency and generalized extensivity in many-body theory. *Mol. Phys.* **103** (2005) 2277.

[33] J. Pipek and P.G. Mazay, A fast intrinsic localization procedure for ab initio and semiempirical linear combination of atomic orbital wave functions. *J. Chem. Phys.* **90** (1989) 4919.

[34] K. Raghavachari, G.W. Trucks, J.A. Pople and M. Head-Gordon, A fifth-order perturbation comparison of electronic correlation theories. *Chem. Phys. Lett.* **157** (1989) 479.

[35] M. Reed and B. Simon, *Methods of Modern Mathematical Physics IV – Analysis of operators*. Academic Press (1978).

[36] T. Rohwedder, *An analysis for some methods and algorithms of Quantum Chemistry*, TU Berlin, Ph.D. thesis (2010). Available on `http://opus.kobv.de/tuberlin/volltexte/2010/2852/`.

[37] T. Rohwedder and R. Schneider, An analysis for the DIIS acceleration method used in quantum chemistry calculations. *J. Math. Chem.* **49** (2011) 1889–1914.

[38] T. Rohwedder and R. Schneider, Error estimates for the Coupled Cluster method. on Preprint submitted to *ESAIM: M2AN* (2011). Available on `http://www.dfg-spp1324.de/download/preprints/preprint098.pdf`.

[39] W. Rudin, *Functional Analysis*. Tat McGraw & Hill Publishing Company, New Delhi (1979).

[40] R. Schneider, Analysis of the projected Coupled Cluster method in electronic structure calculation, *Numer. Math.* **113** (2009) 433.

[41] M. Schütz and H.-J. Werner, Low-order scaling local correlation methods. IV. Linear scaling coupled cluster (LCCSD). *J. Chem. Phys.* **114** (2000) 661.

[42] B. Simon, Schrödinger operators in the 20th century. *J. Math. Phys.* **41** (2000) 3523.

[43] A. Szabo and N.S. Ostlund, *Modern Quantum Chemistry*. Dover Publications Inc. (1992).

[44] G. Teschl, Mathematical methods in quantum mechanics with applications to Schrödinger operators. *AMS Graduate Stud. Math.* **99** (2009).

[45] D.J. Thouless, Stability conditions and nuclear rotations in the Hartree-Fock theory. *Nucl. Phys.* **21** (1960) 225.

[46] J. Weidmann, Lineare Operatoren in Hilberträumen, *Teil I: Grundlagen*, Vieweg u. Teubner (2000).

[47] J. Weidmann, Lineare Operatoren in Hilberträumen, *Teil II: Anwendungen*, Vieweg u. Teubner (2003).

[48] H. Yserentant, Regularity and Approximability of Electronic Wave Functions. Springer-Verlag. *Lect. Notes Math. Ser.* **53** (2010).