Localized states in transition metal oxides
Hozoi, Liviu

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2003

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.
Chapter 2

Theoretical framework

2.1 The embedded cluster model

Within the quantum chemical approach the cluster method is directed to solving the Schrödinger equation for a small but relevant part of a larger system. Certain properties of crystalline solids, e.g., effects connected with the existence of isolated defects and impurities in an otherwise perfect infinite lattice, molecule–surface interactions, core level excitations, localized electronic states in transition metal and rare earth materials etc., are well suited to investigation by cluster methods. The other choice is the periodic approach. It can be fruitful to use a model system in which a unit cell is defined and repeated periodically, so that band structure calculations can be applied to determine specific properties. However, in describing defects, impurities or adsorbates at low-density, the mutual interaction of "defect" sites in neighboring cells limits the accuracy of the results in a periodic approach. Enlarging the cell will reduce such artifacts, but the increase in the computational effort often makes the problem intractable. Also, when combined with advanced quantum chemical wave-function based calculations, the cluster model enables a systematic study of local correlations. An accurate treatment of electron correlation effects is important in 3$d$ and 4$f$ systems.

To insure the proper connection with the rest of the solid various embedding schemes have been developed. Formally, embedding should account for the electrostatic interactions with the surrounding and reflect the fact that the cluster wave-function can extend across the cluster boundaries [1]. The self-consistent solution of the problem is usually based on Green function techniques. This allows to represent the system as a perfect host crystal plus a local perturbation. This "perturbation" may be, for example, induced by a defect or an impurity, related to a photoabsorption process etc. The link between the perfect crystal and the perturbed system is provided by the Dyson equation (for a detailed discussion, see for example ref. [1]). Under the assumption that the perturbation is confined to a compact region in space or, equivalently, to a finite subset of the basis states used to describe the solid, the problem can be reduced to solving the Schrödinger equation within this subdomain alone with an additional energy-dependent embedding potential deduced
CHAPTER 2

from perfect-lattice quantities [2, 3, 4]. These perfect-lattice quantities are calculated using band structure methods. The embedding potential can be regarded as a boundary condition constraining the cluster wave-function to match onto the bulk [2] or as a matrix operator acting on states at the edge of the cluster [4]. Different ways of deriving the working equations are described in ref. [5].

The assumption that the perturbation to the one-electron potential by a defect or impurity is negligible beyond the cluster region provides a lower bound for the size of the cluster. Experience shows that for moderately large clusters the approximation above is reasonable in metals, where screening is strong, and for neutral defects in insulators. For charged defects in insulators and semiconductors such an approximation is not valid, because the unscreened Coulomb potential extends deep into the bulk. For this reason, in insulators, simpler cluster calculations where only the long-range electrostatic field is included constitute a more practical alternative. For many ionic systems such models can be quite accurate because the electronic states are reasonably well localized. In cases with a greater degree of covalency, where the valence electrons are spread more over bonds between neighboring atoms, so-called termination schemes are used to saturate dangling bonds at the boundary of the cluster, for example by tying off these bonds with hydrogen atoms. The advantages of these approaches over formal embedding schemes are conceptual simplicity and wide applicability. Also, possible artifacts of the cluster model can be tested for by varying the cluster size.

In studying the electronic properties of various insulating transition metal oxides we adopt here such an approach. In the applications discussed in the next chapters of the thesis the embedding consists usually of two regions. The first region includes a number of neighboring ions, the nearest and sometimes the next-nearest neighbors, modeled by some "frozen" effective one-electron potentials. These effective operators account for short-range Pauli and exchange interactions between the cluster electrons and electrons at sites nearby, plus electrostatic effects. Normally formal valence states are assigned to these neighboring sites. The second region corresponds to an assembly of point charges extending over few unit cells: formal charges at the lattice positions plus, at the boundaries of this array, a number of charges chosen\footnote{These external point charges can be obtained via a fitting procedure or by using the Evjen method [6].} to reproduce the crystal Madelung potential in the cluster region.

In the present work, calculations are performed on clusters containing one or more transition metal sites plus the adjacent anions. In this case the use of effective potentials is recommended especially in modeling the positive ions next to the cluster, because a simple point charge representation results in an artificial polarization of the ligand orbitals towards these positive charges [7, 8]. For describing the neighboring metal ions we employed either the \textit{ab initio} model potential (AIMP) method developed by Barandiarán and Seijo [8, 9, 10] or so-called nodeless pseudopotentials (PP) as proposed by Durand and Barthelat [46] or Hay and Wadt [12] (for reviews on effective core and embedding potentials, see for example...
High-energy excitations, as occur for example in core level x-ray spectroscopies, and inter-atomic valence charge transfer excitations involve charge redistribution in the cluster region. The polarization of the surrounding crystal modifies the potential seen by the cluster, and in principle a self-consistent treatment is required. In ionic solids one can adopt a discrete representation of the polarizable environment and embed the quantum-mechanical cluster in a set of polarizable entities. One possible approach is the direct reaction field (DRF) method [14, 15, 16], which solves the self-consistency problem for a finite number of polarizabilities. The DRF model was applied mainly to modeling solute–solvent interactions. It was used also to estimate long-range polarization effects for ionization and electron addition processes in nickel oxide [17, 18]. By combining quantum chemical calculations on [NiO$_6$] clusters and the DRF model for describing bulk polarization, the correction to the 3$d$ ionization energy at a Ni$^{2+}$ site was found to be 3–4 eV. In this work ionic polarizabilities were used deduced from the Clausius–Mosotti relation for a rock-salt type lattice.

A second class of methods uses the shell model for treating the long-range lattice polarization. In the shell model each ion in the embedding medium is represented as a point charge core to which an outer spherical shell of $n$ electrons is coupled via a harmonic restoring force with spring constant $k$ [19]. In an electric field the shell retains its spherical charge distribution but moves bodily with respect to the core. The (static) ionic polarizability is then given by $\alpha = (ne)^2/k$, where the $n$ and $k$ parameters are obtained by fitting calculated macroscopic crystal constants to the experimental values. The shell model is implemented for example in the ICECAP (ionic crystal with electronic cluster, automatic program) code [20], which in addition to long-range Coulomb and polarization effects accounts also for the short-range repulsion, via an empirical Buckingham-type potential, and, in defect and impurity problems, for structural readjustments within and around the quantum-mechanical cluster. The ICECAP methodology was intensively applied for studying the electronic structure of transition metal impurities in ionic crystals, $F$ centers, excitons etc., see [21] and references therein.

### 2.2 Wave-function based electronic structure calculations

In this thesis we investigate the local electronic structure of transition metal ions in a number of complex (insulating) oxides, by wave-function based embedded cluster calculations. This implies identifying the electronic ground-state\(^2\), analysis of covalency effects, interpretation of various physical properties, e. g. magnetic couplings, optical spectra, core level excitations etc. The $d$ open-shell structure requires in most of the cases the use of a multiconfigurational approach. The multiconfigurational treatment should consider two

\(^2\)In chapter 5, we propose a new model for the ground-state of the ladder material NaV$_2$O$_5$. 
aspects: first, the presence of several configuration state functions within the $d^n$ manifold and second, the non-negligible contribution from electronic configurations that involve ligand $p$ to metal $d$ charge transfer (CT).

In quantum chemistry, correlation effects are commonly divided into static and dynamic electron correlation\(^3\). Dynamic correlation is due to the interaction between two electrons at short inter-electronic distance. Static, or near-degeneracy, correlation is related to the appearance and strong mixing of two or more electronic configurations having the same, or nearly the same, energy. The multiconfiguration approach mentioned above deals primarily with such near-degeneracy effects. Methods used for recovering dynamic correlation include large scale configuration interaction (CI), perturbation theory (PT), and coupled cluster (CC) theory (for a detailed treatise, see [22] or [23]). The methods employed in this thesis for incorporating static (near-degeneracy) and dynamic correlation effects are discussed in the following subsections.

2.2.1 CASSCF / RASSCF

The electronic wave-function must often comprise several electronic configurations in order to be even qualitatively correct. The natural extension of the single-configuration Hartree–Fock method [24, 25, 26] to systems where strong configurational mixing occurs is the multiconfiguration self-consistent field (MCSCF) model [27]. The most common approach currently is the complete active space SCF (CASSCF) scheme [28], where the user selects a set of chemically important "active" orbitals and the MCSCF configurations are obtained by a full CI within that active space.

In MCSCF theory, the wave-function is written as a linear combination of Slater determinants $|i\rangle$ or configuration state functions $|m\rangle$:

$$|C\rangle = \sum_i C_i |i\rangle = \sum_m C_m |m\rangle.$$ 

The configuration state functions (CSFs) $|m\rangle$ are spin- (and symmetry-) adapted combinations of Slater determinants, i.e. eigenfunctions of the operators for the projected and total spins [29]. The Slater determinants are constructed from a set of real and orthonormal spin orbitals $\{\phi_p(\mathbf{r}, \sigma)\}$, where $\mathbf{r}$ and $\sigma$ are the spatial and spin coordinates, respectively.

In determining the MCSCF wave-function $|C\rangle$, the orbitals are variationally optimized simultaneously with the coefficients of the configuration state functions. The orbitals employed for expressing the wave-function are thus the optimal orbitals for the state at hand and do not introduce a bias towards a particular configuration. The variations of the orbitals and of the CI coefficients can be considered as rotations in an orthonormalized vector space. For example, variations of the orbitals correspond to a unitary transformation of the original orbitals into a new set

$$\tilde{\phi}_p = \sum_q \phi_q U_{qp}. $$

\(^3\)There is, however, no clear-cut partitioning of static and dynamical correlation effects.
The unitary matrix $U$ (i.e. $U^\dagger U = 1$, where $U^\dagger$ is the adjoint of $U$, obtained by transposition followed by complex conjugation) may be written in terms of an anti-Hermitian matrix $\kappa$ as \[22\]:

$$U = \exp(-\kappa), \quad \kappa^\dagger = -\kappa,$$

where the minus sign in the exponential is conventional. When the orthonormal orbitals from which the MCSCF state $|C\rangle$ is constructed are transformed into a different set of orthonormal orbitals, then we obtain a new state, $|\tilde{C}\rangle$. Using the language of second quantization and the notations from ref. \[22\], the new state generated by the unitary transformation $U$ can be expressed in terms of the untransformed, initial state as:

$$|\tilde{C}\rangle = \exp(-\hat{\kappa}) |C\rangle.$$

We have introduced here the anti-Hermitian one-electron operator

$$\hat{\kappa} = \sum_{p,q} \kappa_{pq} a_p^\dagger a_q, \quad \hat{\kappa}^\dagger = -\hat{\kappa},$$

where the summation is over all pairs of creation and annihilation operators $(a_p^\dagger a_q)$ and the parameters $\kappa_{pq}$ are the elements of the anti-Hermitian matrix $\kappa$. So far we have considered an arbitrary unitary transformation of the spin orbitals. In practice we are concerned only with real rotations that preserve spin. We may then write $\hat{\kappa}$ in the form

$$\hat{\kappa} = \sum_{p>q} \kappa_{pq} E_{pq}^-,$$

where the operators $E_{pq}^-$ are antisymmetric combinations of excitation operators

$$E_{pq}^- = E_{pq} - E_{qp}, \quad E_{pq} = a_p^\dagger a_q + a_q^\dagger a_p.$$

The elements $\kappa_{pq}$ $(p > q)$ constitute a set of real parameters that can be considered as the elements of a vector $\mathbf{k}$.

The optimization of the CSF expansion coefficients can be performed by using as variational space the orthogonal complement of a MCSCF initial, reference state $|0\rangle$. This is carried out by means of a configuration vector orthogonal to $|0\rangle$:

$$|C\rangle = \exp(-\hat{\kappa}) \frac{|0\rangle + \hat{P}|c\rangle}{\sqrt{1 + \langle c|\hat{P}|c\rangle}}.$$

The reference state

$$|0\rangle = \sum_m C_m^{(0)} |m\rangle$$
represents our current approximation to the electronic state and is normalized to unity. The state $|c\rangle$ is expanded in terms of the same set of basis states $|m\rangle$ as $|0\rangle$ and contains the free parameters $\{c_m\}$

$$|c\rangle = \sum_m c_m |m\rangle .$$

$\hat{P}$ is the projection operator: $\hat{P} = 1 - \hat{O} = 1 - |0\rangle\langle 0|$. The parameter set $\{c_m\}$ is considered as a vector $c$. The MCSCF wave-function $|C(c, k)\rangle$ is obtained by minimizing the expectation value of the energy with respect to the variational parameters $c$ and $k$:

$$E = \min_{(c,k)} \frac{\langle C|\hat{H}|C\rangle}{\langle C|C\rangle} .$$

In the Born–Oppenheimer approximation and in the absence of external fields, the second-quantization nonrelativistic and spin-free electronic Hamiltonian is given by

$$\hat{H} = \sum_{pq} h_{pq} E_{pq} + \frac{1}{2} \sum_{pqr} g_{pqr} e_{pqrs} + h_{\text{nuc}} .$$

The one- and two-electron matrix elements $h_{pq}$ and $g_{pqr}$ are in the orbital basis $\{\phi_p(r)\}$,

$$h_{pq} = \int \phi^*_p(r) \left( -\frac{1}{2} \nabla^2 - \sum_I \frac{Z_I}{r_I} \right) \phi_q(r) \, dr ,$$

$$g_{pqr} = \int \int \frac{\phi^*_p(r_1) \phi^*_r(r_2) \phi_q(r_1) \phi_s(r_2)}{r_{12}} \, dr_1 dr_2 ,$$

and the scalar term $h_{\text{nuc}}$ represents the nuclear repulsion energy

$$h_{\text{nuc}} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{R_{IJ}} .$$

Here the $Z_I$ are the nuclear charges, $r_I$ the electron–nuclear separations, $r_{12}$ the electron–electron separation, and $R_{IJ}$ the inter-nuclear distances. The summations are over all nuclei. $e_{pqrs}$ is the two-electron excitation operator:

$$e_{pqrs} = E_{pq} E_{rs} - \delta_{qr} E_{ps} = \sum_{\sigma,\tau} a_{p\sigma} a_{r\tau}^\dagger a_{s\tau} a_{q\sigma} .$$

Most of the methods employed for the optimization of the multiconfigurational wavefunction are based on or can be treated as modifications of the second-order Newton–Raphson approach [22, 23]. Another approach is the super-CI method [28, 23], but this will not be discussed here. In the Newton–Raphson scheme, the energy is expanded as a Taylor series to second order in the variational parameters. We arrange the CSF expansion parameters $c$ and the orbital parameters $k$ as a column vector.
\[ \lambda = \begin{pmatrix} c \\ k \end{pmatrix} \]

and in terms of these parameters the second-order MCSCF energy is given by:

\[ E(\lambda) = E(0) + \mathbf{E}^{(1)T} \lambda + \frac{1}{2} \lambda^T \mathbf{E}^{(2)} \lambda. \]

Here \( E(0) \) is the MCSCF energy at the expansion point \( E(0) = \langle 0 | \hat{H} | 0 \rangle \); \( \mathbf{E}^{(1)} \) and \( \mathbf{E}^{(2)} \) are the electronic gradient vector and the Hessian matrix, respectively, at this point, with

\[ E_i^{(1)} = \left( \frac{\partial E}{\partial \lambda_i} \right)_{\lambda=0} \quad \text{and} \quad E_{ij}^{(2)} = \left( \frac{\partial^2 E}{\partial \lambda_i \partial \lambda_j} \right)_{\lambda=0}. \]

The stationary points on the energy surface are determined as solutions to the equations \( \partial E/\partial \lambda_i = 0 \). Setting the derivatives of \( E(\lambda) \) equal to zero leads to the system of linear equations:

\[ \mathbf{E}^{(2)} \lambda = -\mathbf{E}^{(1)}. \]

A sequence of Newton–Raphson iterations is obtained by solving these equations, redefining the solution as the new zero point, recalculating \( \mathbf{E}^{(1)} \) and \( \mathbf{E}^{(2)} \), and returning to the equations above. A detailed discussion of the practical aspects of the optimization of the MCSCF wave-function can be found in refs. [22, 23].

As already mentioned, the most widely used MCSCF method is the CASSCF method [28]. The CASSCF approach involves a partitioning of the occupied orbitals into two sets: inactive, orbitals that are doubly occupied, and active, with occupation numbers between 0 and 2. The remaining orbitals, the secondary or virtual orbitals, are always unoccupied in the wave-function. The CAS space consists of all configurations that can be constructed by distributing the active electrons, i.e., those that are not in the doubly occupied, inactive orbital set, among the active orbitals.

An extension of the CAS model is the restricted active space (RAS) concept [30, 31]. As for CAS wave-functions, the orbitals are divided into three classes: inactive, active and secondary. In addition, there is a further subdivision of the active orbitals: the RAS1, RAS2, and RAS3 categories. In the construction of the RAS wave-function, restrictions are imposed on the number of electrons or holes in the RAS1 and RAS3 subspaces, i.e., a maximum number of holes may be created in RAS1 and a maximum number of electrons is allowed in RAS3; no constraints are placed on the orbital occupations in RAS2. This has the advantage over the CASSCF method that in general more active orbitals can be considered. A variety of wave-functions can be actually created within the RAS framework. For example, a multireference singles and doubles CI (MR-SDCI) wave-function with a CAS reference space may be obtained when RAS1 has at most two holes and RAS3 at most two electrons. The RAS concept combines thus the features of the CAS wave-functions with those of more advanced CI wave-functions where dynamical correlation effects are included.
2.2.2 CASPT2

The treatment of dynamical electron correlation with the multiconfigurational reference function is made by using either the multireference CI (MRCI) method or many-body perturbation theory (MBPT); for reviews, see [22, 23]. Within the MRCI approach, the wave-function is approximated as an expansion in all excited configurations, up to a given level, with respect to a set of reference configurations. Commonly, all single and double excitations from the reference space are included; this is denoted as multireference singles and doubles CI (MR-SDCI). (MR)CI calculations provide a systematic method for treating electron correlation effects. It allows one to approach the exact solution following an order-by-order expansion of the wave-function, i.e. singly excited, doubly excited, triply excited, and so on. Perturbation theory provides an alternative systematic approach to recover the correlation energy. Its advantage over (MR)CI, it is computationally less expensive.

Second-order MBPT is the simplest possible treatment of dynamical electron correlation. A second-order perturbation approach to dynamical correlation, with a reference state given by a multiconfiguration CASSCF-type wave-function has been proposed by Andersson et al. [32, 33] and is referred to as complete active space second-order perturbation theory (CASPT2). In this formulation, the zero-order Hamiltonian is a sum of Fock-type one-electron operators, such that it has $|0\rangle$, the multiconfigurational CASSCF wave-function, as an eigenfunction:

$$\hat{H}_0 = \hat{P}_0 \hat{f} \hat{P}_0 + \hat{P} \hat{f} \hat{P},$$

where $\hat{P}_0 = |0\rangle\langle 0|$ is a projection operator onto the reference function, $\hat{P}$ is a corresponding projection operator for the rest of the configuration space, and $\hat{f}$ is the CASSCF Fock operator:

$$\hat{f} = \sum_{p,q} f_{pq} E_{pq} = \frac{1}{2} \sum_{p,q} \sum_{\sigma} \langle 0 | [a_{q\sigma}^\dagger, [a_{p\sigma}, \hat{H}]] + |0\rangle E_{pq}.$$  

The CI space is partitioned into four subspaces: 0 – the reference function, $K$ – the rest of the CASCI space, $SD$ – all singly and doubly excited configuration state functions with respect to the CAS reference, and $X$ – the rest of the CI space. This is accomplished by introducing the orthogonal projectors $\hat{P}_K + \hat{P}_{SD} + \hat{P}_X = \hat{P}$, with $\hat{P}_0 + \hat{P} = 1$. The block diagonal structure imposed by the projectors insures that the corrections to the wave-function truncate at a finite level of excitation.

The perturbation part of the Hamiltonian is the difference between the full Hamiltonian and $\hat{H}_0$. Using the so-called internally contracted scheme, we write the first-order wave-function correction as a linear combination of all single and double excitations from $|0\rangle$:

$$|0^{(1)}\rangle = \sum_{p,q} C_{pq}^{(1)} E_{pq} |0\rangle + \sum_{p,q,r,s} C_{pqrs}^{(1)} e_{pqrs} |0\rangle.$$  

In the summations, $q$ and $s$ are occupied, inactive or active, $p$ and $r$ are noninactive, i.e. active or secondary. All single and double excitations with respect to $|0\rangle$ are included in
the wave-function, except those which have all four indices in the active space. Such terms belong to the space \( K \) and do not interact with \( |0\rangle \). The internally contracted formalism \([34]\) reduces considerably the number of terms in the perturbation series, at the cost of greater complexity in the individual terms.

The first- and second-order CASPT energy corrections are calculated as \([32, 33]\):

\[
E^{(1)} = C^{(0)T} U C^{(0)} \quad \text{and} \quad E^{(2)} = C^{(1)T} U C^{(0)},
\]

where \( C^{(0)} \) defines the zero-order state and we used the partitioning of the Hamiltonian matrix \( H = H_0 + U \). The first-order wave-function \( C^{(1)} \) is obtained as an iterative solution \([32, 33]\) to a set of linear equations of the type \([22]\):

\[
C^{(1)} = -P (H_0 - E^{(0)} S)^{-1} P U C^{(0)},
\]

where \( P = 1 - C^{(0)} C^{(0)T} \), \( S_{ij} = \langle i | j \rangle \), and \( |i\rangle \) refers to one of the contracted configurations \( E_{pq} |0\rangle \) or \( e_{pqr} |0\rangle \).

Within the embedded cluster approach, CASSCF/CASPT2 calculations have been successfully employed for the interpretation of optical absorption spectra in several transition metal oxides \([36, 37]\) and for the investigation of magnetic interactions, see section 2.3 of the thesis.

The extensive analysis by de Graaf et al. \([36, 37]\) demonstrates that CASSCF/CASPT2 is an efficient method for studying low-energy electronic excitations in insulating (3d) transition metal oxides, metal \( d \rightarrow d \) as well as ligand \( p \) to metal \( d \) CT excitations. de Graaf et al. \([36, 37]\) showed that three effects are important in order to obtain an accurate description of such \( d^n \) and \( d^{n+1}L \) states (\( L \) represents a ligand \( p \) hole). In the first place, there is the mixing of the different \( d^n \) CSFs in the wave-function, for which an active space containing the transition metal \( d \) orbitals is sufficient. Secondly, important correlation effects, both static and dynamic, occur among the metal \( 3d \) and ligand \( 2p \) valence electrons. To include these correlation effects, the active space should contain the metal \( 3d \) and ligand \( 2p \) orbitals, plus a set of correlating orbitals of the same type, usually referred to as ”prime” orbitals — metal \( d' \) \([38, 39, 40, 41, 36, 37]\) and ligand \( p' \) \([37]\). Such an active space, \( \text{CAS} - p \ d \ d' p' \), treats a large part of the valence electron correlation in a variational manner, and hence avoids both the overestimation of the \( d - d \) correlation energy that was observed in the CAS – \( d \) PT2 treatment \([40, 41, 37]\) and the occurrence of intruder states \([23, 35]\) connected to charge transfer excitations \([36, 37]\). However, due to the limitations of the current computer resources, in most of the cases only a small number of ligand \( p, p' \) orbitals can be included in the CAS \([36, 37]\). In the third place, we need to include the (semi)core–valence correlation effects due to the metal \( 3s, 3p, 3d \) and ligand \( 2s, 2p \) electrons. Second-order perturbation theory is able to provide a quite accurate treatment of these last effects, see for example \([40, 41, 36, 37]\).
2.2.3 Non-orthogonal CI / State Interaction

Usually several electronic states, ground-state and excited states, are of interest. In many cases different states of interest may not be described compactly with one common orbital basis. Different electronic states have often very differently shaped orbitals. The error introduced by applying a common set of orbitals can then be recovered only by using extended CI, or MRCI, expansions.

However, if separate SCF calculations are performed for different states of the same spin and symmetry, the resulting wave-functions will not be orthogonal to each other. To arrive at a more convenient description, we may define new wave-functions as optimized linear combinations of the non-orthogonal, interacting wave-functions. The optimization of the coefficients $C_m$ of a wave-function $|C\rangle$ expanded in mutually non-orthogonal many-electron functions $|m\rangle$,

$$|C\rangle = \sum_m C_m |m\rangle,$$

leads to secular equations with the non-orthogonal wave-functions as a many-electron basis. In cases where $|m\rangle$ are single CSFs, this approach is called non-orthogonal CI (NOCI) [42, 43, 44, 45, 46]. In other cases, for example when each of a number of states is approximated with a MCSCF wave-function, the many-electron basis functions $|m\rangle$ may consist of more than one CSF. If this is the case, it is more appropriate to denote the approach state interaction (SI). The term state interaction has been used mainly for solving the secular equations in terms of individually optimized CASSCF or RASSCF wave-functions [47, 48].

In both NOCI and SI, the wave-functions are thus expanded in terms of “relaxed” many-electron functions that are each expressed in terms of their "own" optimized orbital set. Such an approach can yield more transparent wave-functions with a much shorter expansion length. Also, individually optimized wave-functions are important for the accurate calculation of transition properties like transition moments and oscillator strengths, see for example [42, 48] and [23], sections 34.3–34.4. As shown in [49] for molecular systems, the separate optimization of the orbitals is more important for obtaining accurate values of the transition moments than extensive inclusion of dynamical correlation.

The computation of Hamilton matrix elements between mutually non-orthogonal (MC) SCF wave-functions is not trivial. However, efficient schemes have been developed, to calculate such matrix elements between both non-orthogonal individual CSFs [43, 45] and non-orthogonal multiconfigurational RASSCF wave-functions [47, 48].

The NOCI/RASSI approach is well suited to study systems for which large (differential) energetic effects are obtained with a broken-symmetry (MC)SCF wave-function. Symmetry breaking can occur in the SCF description of hole states in symmetrical systems, see for example [43, 44, 45, 46]. The energy lowering caused by symmetry breaking of SCF hole-state wave-functions is due to the relaxation energy. The main effect is an atomic effect: if an electron hole is delocalized over $n$ symmetry-equivalent sites only about $1/n$ of the atomic relaxation energy is obtained. Only if a hole is created on one site the full relaxation energy is obtained [50]. An example of this situation is discussed in the third chapter of
the thesis, in connection with the investigation of metal 3s x-ray photoelectron excited states in nickel oxide and manganese oxide.

2.3 Magnetic interactions

Interactions between localized spin moments in transition metal materials are the source of remarkable phenomena. Examples are the large diversity of magnetically ordered structures in these compounds, the spin-gap formation in some low-dimensional copper and vanadium oxides (SrCu$_2$O$_3$ [18], CaV$_2$O$_5$ [10], and NaV$_2$O$_5$ [53]), the spin-Peierls dimerization in CuGeO$_3$ [54] etc.

Most of the (undoped) transition metal oxides, including the ones mentioned above, exhibit insulating properties and antiferromagnetic (AFM) order. Ferromagnetic (FM) interactions occur only in a few oxides, e.g. LaMnO$_3$, Fe$_3$O$_4$, or CrO$_2$, and are often coupled with metallic-like behavior. The nature of magnetic interactions is best understood in insulators, where the spins are well localized, though the theory is still complex. Especially in insulators, many of the magnetic properties can be described in terms of effective Heisenberg Hamiltonian models [55] with predominant nearest-neighbor $< i, j >$ interactions:

$$\hat{H} = - \sum_{< i, j >} J_{ij} \mathbf{S}_i \cdot \mathbf{S}_j .$$

This expression is the generalization of the spin Hamiltonian for a two-site system. For only one spin at each site ($S_1 = S_2 = 1/2$), the derivation of the relation that connects the Heisenberg coupling parameter $J$ to the singlet–triplet splitting $E_S - E_T$ was given in ref. [56]. Since in this case the operator

$$\mathbf{S}_1 \cdot \mathbf{S}_2 = 1/2 (\mathbf{S}^2 - S_1^2 - S_2^2),$$

where $\mathbf{S} = \mathbf{S}_1 + \mathbf{S}_2$, has eigenvalues $-3/4$ in the singlet ($S = 0$) state and $1/4$ in the triplet ($S = 1$) state, one can define an operator

$$\hat{H} = 1/4 (E_S + 3 E_T) - (E_S - E_T) \mathbf{S}_1 \cdot \mathbf{S}_2$$

that has eigenvalue $E_S$ in the singlet state and $E_T$ in the triplet state. The first term is of no interest to us, and we may write the spin Hamiltonian as

$$\hat{H} = - J \mathbf{S}_1 \cdot \mathbf{S}_2 , \quad J = E_S - E_T .$$

This implies that for ”parallel” (ferromagnetic) spin alignment $J$ is positive ($E_T < E_S$), whereas for ”antiparallel” (antiferromagnetic) spins $J$ is negative ($E_S < E_T$).

The question of when the nearest-neighbor Heisenberg Hamiltonian model is justified is a complex matter. The validity of this approximation has been discussed in refs. [57, 58, 59]. There are studies indicating that higher order terms, e.g. biquadratic contributions
CHAPTER 2.

\( (S_i \cdot S_j)^2 \), and/or next-nearest-neighbor couplings are actually needed for explaining the magnetic properties of certain materials. However, even starting from the simple Heisenberg model with couplings only between nearest neighbors, the problems encountered in trying to deduce the behavior of various magnetically ordered systems are quite difficult [60].

The antiferromagnetic couplings in transition metal insulating compounds are qualitatively explained by the Anderson "delocalization" or "kinetic" (super)exchange mechanism [61]. The term superexchange was originally introduced by Kramers [62] in 1934, in an early attempt to explain the magnetic interaction taking place between spin moments that are fairly well separated by diamagnetic groups. However, it is the subsequent work of Anderson which provides the basic ingredients for understanding antiferromagnetism. Neglecting some smaller terms, the superexchange interaction can be viewed as the result of two effects: direct exchange, due to the usual Coulomb interaction, and kinetic exchange, arising from electron transfer from one magnetic center to the other [61]. The first term is always ferromagnetic. The second contribution is of opposite sign and usually more important than the direct exchange, favoring AFM order. Exceptions occur in systems with orbital ordering, as in LaMnO_3, where virtual inter-site excitations may result in a FM alignment.

Over the last years ab initio quantum chemical calculations applied to embedded cluster models proved to be valuable tools not only to accurately predict the magnitude of the exchange coupling parameters, but also to elucidate the underlying microscopic mechanisms. Within such an approach, the coupling constants are derived from the energies of the states arising from different spin configurations. Expressions similar to the relation given above for a model system with two magnetic centers and one spin moment at each site can be obtained for systems with more than one unpaired electron per magnetic center and/or more than two centers, by establishing a correspondence between the energy expectation values of the N-electron cluster Hamiltonian and the eigenvalues of the Heisenberg Hamiltonian. For example, for a Ni\(^{2+}\)-O\(^{2-}\)-Ni\(^{2+}\) unit in nickel oxide, with two unpaired electrons at each Ni site \((S_1 = S_2 = 1)\), the eigenvalues of the Heisenberg Hamiltonian

\[
\hat{H} = -J \mathbf{S}_1 \cdot \mathbf{S}_2 = -J \left[ (s_{11} + s_{12}) \cdot (s_{21} + s_{22}) \right]
\]

for the quintet, triplet, and singlet spin functions of highest \(M_s\) value,

\[
|Q, M_s=2\rangle = \alpha\alpha\alpha\alpha ,
\]

\[
|T, M_s=1\rangle = \frac{1}{2} (\beta\alpha\alpha\alpha + \alpha\beta\alpha\alpha - \alpha\alpha\beta\alpha - \alpha\alpha\alpha\beta) ,
\]

\[
|S, M_s=0\rangle = \frac{1}{2\sqrt{3}} (2\beta\beta\alpha\alpha + 2\alpha\alpha\beta\beta - \beta\alpha\beta\alpha - \beta\alpha\alpha\beta - \beta\beta\beta\alpha - \beta\beta\alpha\beta) ,
\]

are \(-J\), \(J\), and \(2J\), respectively. The one-to-one mapping between these quantities and the energy expectation values \(E_Q\), \(E_T\), and \(E_S\) of the full Hamiltonian provides a way
to estimate the strength of the magnetic interaction: $E_Q - E_T = -2J$, $E_T - E_S = -J$ [63, 37]. If $J$ may be considered as describing a genuine two-center interaction, it may be extracted from calculations of the electronic states of material models containing two magnetic centers only. From the explicit investigations of Illas et al. [63] and de Graaf et al. [64], where the nearest-neighbor $J$ value obtained from two-center clusters was the same as obtained when more magnetic centers were included in the model, it appears that this is often the case, at least in compounds like KNiF$_3$, K$_2$NiF$_4$, NiO, La$_2$CuO$_4$, and Li$_2$CuO$_2$.

In a wave-function based embedded cluster approach, the Anderson kinetic exchange can be accounted for by complete active space configuration interaction (CASCI) schemes where the active space includes the open-shell orbitals of the transition metal ions, the "magnetic" orbitals. Such an active space contains the leading ground-state configuration, $d^{n-}L-d^{n}$, and all metal to metal symmetry-allowed charge transfer configurations of the type $d^{n-i}L-d^{n+i}$, where $L$ represents the bridging ligand(s). Calculations at this level of accuracy usually reproduce the sign of the interaction, but the magnitude of the coupling parameter is severely underestimated. Several methods have been developed in order to include effects that go beyond the CASCI approximation. By gradually improving the description of the cluster wave-function, detailed information can be obtained about the physical mechanisms determining $J$.

A procedure successfully applied for the study of magnetic interactions is the difference dedicated configuration interaction (DDCI) [65, 66]. The states defining $J$ are expressed in terms of CI expansions including single and double excitations from a CAS formed by the open-shell orbitals. The CI list is, however, restricted by eliminating all the double excitations from inactive to virtual orbitals, i.e. those that do not involve any of the open-shell orbitals. For a common set of orbitals, up to second order, these determinants only cause a shift in the diagonal matrix elements of the CI-matrix and hence do not contribute to the energy difference of the states of interest [66, 67]. Using the labels from ref. [66], $(p, q, ...)$ for the inactive orbitals, $(a, b, ...)$ for the active orbitals, and $(i, j, ...)$ for the virtuals, such excitations are denoted as $pq \rightarrow ij$. By excluding these doubly excited determinants from the CI list, the size of the problem is reduced considerably. It has been shown that this computational scheme predicts spin coupling parameters in very good agreement with the experimental data in several solid state ionic insulators [68, 69] and molecular complexes [70, 71].

DDCI calculations may become computationally too expensive for systems where large clusters and/or large reference active spaces are needed. In such cases, one has to restrict oneself to the so-called DDCI2 scheme [65], where $pa \rightarrow ij$ and $pq \rightarrow cj$ excitations are also excluded from the CI expansion. Physical effects taken into account by this procedure beyond direct and kinetic exchange are the so-called spin polarization and the dynamical polarization of the ligands under the effect of $d^{n-i}L-d^{n+i}$ charge fluctuations, and have been discussed for $S=1/2$ systems in ref. [71]. Although DDCI2 yields normally only 50–80% of the experimental $J$, see [68, 69] and references therein, calculations at this level can be used in a semi-quantitative way for predicting trends when structural distortions
or changes in the chemical composition are introduced in the system.

The analysis of Calzado et al. [71] indicates that the difference between DDCI2 and DDCI is due to a more effective relaxation of the metal to metal and ligand to metal charge transfer configurations. The importance of orbital relaxation in the CT component of the wave-function was originally investigated by van Oosten et al. [72, 73] in several copper oxide systems by means of non-orthogonal CI. These calculations were based on small CI expansions including only O→Cu excitations, where both the reference and the CT excited configurations were expressed in terms of fully optimized orbitals. It has been found that the AFM exchange is strongly enhanced by admixing relaxed charge transfer states, with estimates of the coupling constants in close agreement with the experimental values, although somewhat smaller than the DDCI results.

The NOCI expansion is much shorter than for a conventional CI and offers an insightful physical picture of such CT effects. However, since it has not received numerous applications, further research is needed to check whether its good performance in cuprates is not fortuitous. Also, the optimization of the excited state orbitals is often rather tedious. An approach which avoids both the optimization problems encountered in non-orthogonal CI and the relatively high computational demand of DDCI is the CASPT2 method, already mentioned in a previous section. It has been found that, when based on a minimal CAS including only the open-shell orbitals, CASPT2 usually underestimates the strength of the magnetic interaction [37, 74, 75]. However, rather accurate exchange couplings are obtained when the minimal CAS is extended with those ligand $p$ orbitals that mediate the superexchange plus the correlating $d'$ and $p'$ counterparts. For several transition metal compounds, more than 80% of the experimental estimate is reproduced by such CAS–$p\,d\,d'p'$ calculations, see [37, 74, 75] and also the fifth chapter of this thesis work. This indicates once more that in order to obtain high-quality results, it is essential to properly account for the relaxation of the ligand to metal charge transfer configurations [72, 73].
References


THEORETICAL FRAMEWORK


