Localized basis functions, advantages and disadvantages
Two groups of methods using local basis functions
General projected basis sets
Second quantization in singular nonorthogonal basis
Solving The MP1 equations in the local basis,
Some results
Conclusion
Why use local orbitals in correlated calculations

The Coulomb hole

- Caused by the repulsion of electrons
- Local: only non vanishing when two electrons are close to each other

The Coulomb hole may be described in different ways

- **Non-local** functions → steep rise in complexity
- **Local functions** → linear rise in complexity
Pros and cons of local orbitals

**Advantages**
- Local functions are expanded in a local basis
- Allows the development of linear scaling or low-scaling methods
- Avoids intermediate constructs like molecular orbitals → Formulation directly in computable terms

**Disadvantages**
- Our normal mental picture of molecular orbitals is lost
- No simple division into occupied and virtual orbitals
- Loss of diagonal dominance of matrices → iterative methods may be very slowly convergent

Localized orbitals are usefull when it is more important to distinguish between near and far than between occupied and unoccupied
Two main groups of local correlation methods

The domain approach

- Divide molecule into domains
- Localize occupied orbitals, redundant basis for virtual orbitals
- Use domains to discriminate between strong and weak interactions
- Computationally very efficient, does not go to standard MP2 limit
- Some arbitrariness, assumes that occupied orbitals may be localized
- Pioneered by P. Pulay, important work by H.J. Werner, M. Schütz

The domain-free approach

- Develop local redundant basis for occupied and virtual orbitals
- Use sparse arithmetic and storing to exploit locality
- More expensive, dimension of basis for occupied orbitals = dimension of basis set, does go towards MP2 limit
- Pioneered by M. Head-Gordon and G. Scuseria
General local basis expansions of occupied and virtual orbitals

Goals: Basis functions that are both local and provide diagonal dominant matrices

Procedure

1. Standard atomic orbital basis \( |\chi\rangle \), dimension \( N \), metric \( S_{ij}^\chi = \langle \chi_i | \chi_j \rangle \)
2. Transform to another local basis \( |\phi\rangle = |\chi\rangle Y \)
3. Obtain redundant basis sets of dim \( N \) for occupied and for virtual orbitals by \( |\psi\rangle = (|\psi_k\rangle, |\psi_a\rangle) = |\chi\rangle Z \)
4. Obtain pseudo-biorthonormal basis to simplify equations
Step 2: $|\phi\rangle = |\chi\rangle Y$: local basis of dimension $2N$ allowing convergence of equations

- Orthonormal local basis sets provide diagonal dominant matrices!
  1. Cholesky factorization: $S^x = U^T U$, $Y = U^{-1}$ (Used here !)
  2. Symmetric orthogonalization: $Y = (S^x)^{-\frac{1}{2}}$
- Both may be realized in linear scaling fashion
- $Y = 1$ leads to standard biorthonormal approach

Step 3: $|\psi\rangle = |\chi\rangle Z$: separate basis sets for occ. and virt.

- $Z = (P^x S^x Y, Q^x S^x Y)$
- $P^x$ is density matrix in atomic orbital basis, $Q^x = (S^x)^{-1} - P^x$
- Leads to a nonorthogonal basis of $2N$ functions- singular

Less local than projected AO’s $(|\chi\rangle P^x S^x, |\chi\rangle Q^x S^x)$
The Pseudo-biorthonormal basis

Standard biorthonormal approach: example, biorthonormal atomic orbital basis

- Introduce a basis \( \langle \tilde{\chi} \rangle \) so \( \langle \tilde{\chi}_i \vert \chi_j \rangle = \delta_{ij} \)
- \( \langle \tilde{\chi} \rangle = (S^x)^{-1} \langle \chi \rangle \)
- Assumes a nonsingular metric
The Pseudo-biorthonormal basis

Extension to singular basis

- $S^\psi = \langle \psi | \psi \rangle$ is singular $2N \times 2N$ singular metric

- Introduce $S^{\psi \Theta} = \begin{pmatrix} P^\phi & 0 \\ 0 & Q^\phi \end{pmatrix}$

- Has some of the properties of standard inverses, for example $S^\psi S^{\psi \Theta} S^\psi = S^\psi$

- Is the standard general (Penrose) inverse if $|\phi\rangle$ is an orthonormal basis, else $(S^\psi S^{\psi \Theta})^\dagger \neq S^\psi S^{\psi \Theta}$

- Introduce $\langle \bar{\psi} | = S^{\psi \Theta} \langle \psi |$

- Overlap $P = \langle \bar{\psi} | \psi \rangle = \begin{pmatrix} P^\phi S^\phi & 0 \\ 0 & Q^\phi S^\phi \end{pmatrix}$

- $P$ is a projection matrix, for example $P^2 = P$

- As close as possible to standard biorthonormal basis
Second quantization in a singular nonorthogonal basis

Purpose

Allow the development of methods using SQ directly in the $\psi$ basis without backtransforming from MO basis

<table>
<thead>
<tr>
<th></th>
<th>Standard nonorth.</th>
<th>singular nonorth.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Biorthonormal operator</td>
<td>$\bar{a}<em>P = \sum_R (S)</em>{PR}^{-1} a_R$</td>
<td>$\bar{a}<em>P = \sum_R (S)</em>{PR}^{\psi\Theta} a_R$</td>
</tr>
<tr>
<td>Anticommutation</td>
<td>$[a_P^\dagger, \bar{a}<em>Q]</em>+ = \delta_{PQ}$</td>
<td>$[a_P^\dagger, \bar{a}<em>Q]</em>+ = \cal{P}_{QP}$</td>
</tr>
<tr>
<td>1-elec operator</td>
<td>$\hat{h} = (S^{-1}h)_{PQ} a_P^\dagger \bar{a}_Q$</td>
<td>$\hat{h} = (S^{\psi\Theta}h^\psi)_{PQ} a_P^\dagger \bar{a}_Q$</td>
</tr>
</tbody>
</table>
The MP1 equations in the $\psi$ basis

Inserting the expansion

$$\hat{T}_2 = \frac{1}{2} \sum_{abij} T^{ab}_{ij} \bar{E}_{ai} \bar{E}_{bj}, \quad E_{ai} = \sum_{\sigma} a_{a\sigma}^{\dagger} \bar{a}_{i\sigma} \quad a, b: \text{virtual}, \; i, j: \text{occupied}$$

In the $N^4$ MP1-equations

$$\langle i_j^{ab} | \hat{H} + [\hat{F}, \hat{T}_2] | HF \rangle = 0 \quad (1)$$

gives the $N^4$ equations ($\chi$-indeces neglected)

$$\sum_{cdkl} (QS)_{ac} (QS)_{bd} (PS)_{ki} (PS)_{lj} \left( \bar{g}_{ckdl} + (WT)_{cd}^{kl} \right) = 0$$

$$W = \bar{F} \otimes 1 \otimes 1 \otimes 1 + 1 \otimes \bar{F} \otimes 1 \otimes 1 + 1 \otimes 1 \otimes \bar{F} \otimes 1 + \ldots$$

The $\text{QSQSPSPS}$-term come from the general anticommutation; previously neglected by us and others
The essential equations

\[ \sum_{cdkl} (QS)_{ac} (QS)_{bd} (PS)_{ki} (PS)_{lj} (\tilde{g}_{ckdl} + (WT)^{cd}) = 0 \]

**Standard (Previous approach)**

- Solve \( \tilde{g}_{ckdl} + (WT)^{cd}_{kl} = 0 \)
- Solves in addition to the essential equations also a set of nonessential equations
- The nonessential equations are typically not automatically fulfilled \( \rightarrow \) slower convergence

**Current approach**

- Include projection operators to solve only the essential conditions
- Each iteration becomes more complicated, but (hopefully) faster convergence
Improve convergence by modifying the essential equations

\[ \sum_{cdkl}(QS)_{ac}(QS)_{bd}(PS)_{ki}(PS)_{lj} \left( \bar{g}_{ckdl} + (WT)_{kl}^{cd} \right) = 0 \]

**Change W**

- Any matrix \( M \) may be added to \( W \) without changing the solution if \( QSQSPSPSM = 0 \)
- Does not change solution, but may improve convergence by giving equations that are more easy to precondition
- Exploits that a singular set of linear equations may be modified without modifying the solution

**Possible choice of \( M: W \) with interchange of blocks in \( \bar{F} \)**

<table>
<thead>
<tr>
<th>( \bar{F} ), standard</th>
<th>( \bar{F} ), blocks interchanged</th>
</tr>
</thead>
<tbody>
<tr>
<td>( F = \begin{pmatrix} P\bar{F} &amp; 0 \ 0 &amp; Q\bar{F} \end{pmatrix} )</td>
<td>( A = \begin{pmatrix} Q\bar{F} &amp; 0 \ 0 &amp; P\bar{F} \end{pmatrix} )</td>
</tr>
</tbody>
</table>
Example of convergence of iterative method

$C_6H_2$ in various basis sets, DIIS with diagonal preconditioner

Improvements compared to standard iterative approach

- Use of general local basis $\rightarrow$ convergence
- Focus on the essential equations $\rightarrow$ improved convergence
- $W$ replaced by $W + M$ $\rightarrow$ improved convergence

Number of iterations as function of convergence threshold

<table>
<thead>
<tr>
<th>Basis</th>
<th>Threshold</th>
<th>All eqs.</th>
<th>Ess. eqs.</th>
<th>All eqs.</th>
<th>Ess. eqs.</th>
</tr>
</thead>
<tbody>
<tr>
<td>STO-3G</td>
<td>$10^{-3}$</td>
<td>9</td>
<td>8</td>
<td>11</td>
<td>7</td>
</tr>
<tr>
<td>STO-3G</td>
<td>$10^{-5}$</td>
<td>17</td>
<td>13</td>
<td>24</td>
<td>11</td>
</tr>
<tr>
<td>cc-pVTZ</td>
<td>$10^{-3}$</td>
<td>34</td>
<td>24</td>
<td>24</td>
<td>19</td>
</tr>
<tr>
<td>cc-pVTZ</td>
<td>$10^{-5}$</td>
<td>94</td>
<td>53</td>
<td>62</td>
<td>39</td>
</tr>
</tbody>
</table>

The number of iterations may be reduced by up to a factor of 3
Solution of the linear equations using DIIS

- To solve the linear equations we employ DIIS, rather than PCG.
- The traditional formulation of DIIS depends on the dimension of the subspace, also for linear equations with symmetric matrices.
- Reformulation ensures that three vectors in subspace is mathematically identical to using all vectors in subspace.
- Ingredients: save optimal vectors and use norm containing preconditioner.
Preliminary scalings

$C_nH_2$ for $n = 10, 20, 24, 30, 34$ in the cc-pVDZ basis

Upper figure: CPU time
Far from linear scaling (fourth-order scaling!)

Lower figure: # of integrals > $10^{-8}$
- Two-electron integrals $g$ in the projected basis have significant more non-vanishing elements than in the atomic basis $g^\chi$
- Number of projected integrals $\approx$ cubic scaling
- Number of atomic integrals $\approx$ quadratic scaling
A theoretical framework for working with singular basis sets has been developed.

Transformation to an orthonormal basis before projecting improves convergence significantly.

By solving only the essential equations and modifying the equations, the iterations may be reduced with up to a factor of 3.

Orbitals obtained by symmetric orthogonalization converge somewhat faster and regular than orbitals obtained by Cholesky-factorization.

Major formal and computational restrictions have been eliminated.

A long way to go...
Acknowledgements

Financial support

- Danish Research Council for a research center
- Finnish Cultural Foundation and the Finnish Academy
- The NANOQUANT EU-network
- CSC in Espoo and DCSC for computational resources
- The Lundbeck Foundation

The papers

- O. Christiansen et al, JCP, 124, 084103 (2006): our initial discussion of approach
- B. Jansik et al, JCP, 126, 124104 (2007): $S^{-\frac{1}{2}}$