The development of methods for the calculation of the response of molecular systems to internal and external perturbations

Trygve Helgaker

Filip Pawłowski, Simen Reine, Erik Tellgren, Mark Watson

Centre for Theoretical and Computational Chemistry, University of Oslo, Norway

Stinne Høst, Thomas Kjærgaard, Branislav Jansík, Poul Jørgensen, Jeppe Olsen

University of Aarhus, Denmark

Sonia Coriani, University of Trieste, Italy

Paweł Sałek, Royal Institute of Technology, Sweden

Michael Peach and David Tozer, University of Durham, UK

Odyssey 2008

Mathematical and Computational Aspects of Molecular Electronic Structure Calculations

Campus Saint-Jean, University of Alberta, Edmonton, Alberta, Canada

June 1–5, 2008
Calculation of energies and properties: new techniques suited large systems

• With improvements in computer technology and computational techniques, quantum chemistry is being applied to increasingly large systems
  – Hartree–Fock and Kohn–Sham theory can nowadays be applied to hundreds of atoms

• However, larger systems have always presented a problem to quantum chemistry
  – steep increase in computational cost \( N^4 \) or worse
    * new techniques must be introduced to curb cost (locality)
  – old models may no longer applicable
    * for example, the CI model works poorly for many electrons
    * superseded by the size-extensive coupled-cluster model
  – new physical phenomena become important
    * long-range correlation and exchange may dominate chemistry
    * traditional DFT does not account for these effects

• We shall here consider new methods, better suited for large systems
  – direct optimization of density matrix
  – calculation of properties from density matrix
Standard self-consistent field (SCF) optimization

- Consider the standard Roothaan–Hall (RH) SCF optimization scheme:
  - in an AO basis, make an initial guess of the density matrix $D_0$
  - at each iteration $n$, construct Fock/KS matrix from current density matrix $D_n$:
    \[ F_n = h + G(D_n) \]
  - next, diagonalize the Fock/KS matrix to yield new density matrix $D_{n+1}$:
    \[ F_n C = SC\epsilon, \quad D_{n+1} = C_{occ}C_{occ}^T \]
  - repeat until convergence (self consistency) $D_{n+1} \approx D_n$

- The RH SCF scheme works only in simple cases (oscillations and divergence)

- The DIIS (direct inversion in iterative subspace) method greatly improves convergence
  - the Fock/KS matrix $F_{n+1}$ is not constructed from $D_{n+1}$
  - instead, it is constructed from an averaged density matrix
    \[ \overline{D} = \sum_{i=0}^{n} c_i D_i, \quad \sum_i c_i = 1, \quad F_{n+1} = F(\overline{D}) \]

- The RH/DIIS scheme is the standard optimization scheme in HF/KS theories.
Cost of SCF optimizations in large systems

- Consider the optimization of the SCF energy (here LDA) of molecular systems:

  - small systems dominated by KS-matrix evaluation, with linear scaling
  - large systems dominated by SCF diagonalization, with cubic scaling

- To achieve linear scaling, we must avoid diagonalization (and MOs)!
SCF convergence is typically more difficult in larger systems
- small (or negative) HOMO–LUMO gaps and small Hessian eigenvalues in DFT
- 6-31G HOMO–LUMO gap and lowest Hessian eigenvalue in alanine residue peptides

Large systems therefore force us to reconsider energy optimization
Properties for large systems

- B3LYP/6-31G excitation energies of alanine residue peptides

- the RH/DIIS scheme converges, for sufficiently large systems, to an excited state
- the B3LYP ground state becomes unphysical for large systems

- Why does RH/DIIS not yield a ground state? Why is the ground state so strange?
Density-matrix based optimization

- Traditional SCF theory rests on the concept of molecular orbitals (MOs):
  \[ \mathbf{FC} = \mathbf{SC} \epsilon \rightarrow \text{Roothaan–Hall equations} \]
  - diagonalization is expensive and ill suited to parallelization
  - convergence can be difficult (oscillations, divergence, saddle points)
- Fortunately, MOs are not needed: all information is contained in the density matrix
  \[ \mathbf{D} = \mathbf{C}^{\text{T}} \mathbf{C} \rightarrow \text{density matrix} \]
  - it is sufficient to work in terms of the one-electron density matrix
  - this is particularly advantageous if it can be compactly (sparsely) represented
- This approach is taken in our work, but then some questions arise:
  - Q1: In what basis do we represent the density matrix?
  - Q2: How do we optimize the density matrix, without recourse to orbitals?
  - Q3: How do we calculate its response to perturbations?
- All tasks should be accomplished in terms of elementary matrix manipulations
  - additions, multiplications, trace operations—no diagonalizations
  - with sparsity, linear scaling is then achievable for large systems
  - diagonalization-free methods are simpler to parallelize
Energy minimization

• Consider the direct optimization of the density matrix in the AO basis:

\[ E(D) = \text{Tr} Dh + \frac{1}{2} \text{Tr} DG(D) \]

• There are constraints on the density matrix

\[ D = D^T \quad \text{← the density matrix must be symmetric} \]
\[ \text{Tr} DS = N \quad \text{← it must contain exactly } N \text{ electrons} \]
\[ DSD = D \quad \text{← its occupations numbers must be one or zero} \]

  – any optimization must obey these constraints

• There are different ways of dealing with these constraints

  – perform a constrained minimization (Lagrange multipliers)
  – introduce a penalty for disobeying the constraints
  – work in terms of a McWeeny purified density matrix \( \tilde{D} = 3DSD - 2DSDSD \)

• We shall pursue a different approach

  – introduce a parameterization of \( D \) that obeys the constraints
Exponential parameterization of the density matrix

- In a real, nonorthogonal AO basis, with $S \neq I$, let $D$ be a valid HF/KS matrix:

  $D = D^T$, $\text{Tr} DS = N$, $DSD = D$

  symmetry \hspace{1cm} trace \hspace{1cm} idempotency

- Any other valid density matrix $D(X)$ can then be generated from this matrix:

  $D(X) = \exp(-XS)D \exp(SX)$, $X^T = -X$

  exponential parameterization \hspace{1cm} antisymmetric


- We can obtain any valid density matrix, in the AO basis, without recourse to MOs!
  - in particular, we may optimize the energy by freely varying $X_{\mu \nu}$ with $\mu > \nu$:

    $E_{\text{min}} = \min_X [\text{Tr} D(X)h + \frac{1}{2} \text{Tr} D(X)G(D(X))]$

    - we may apply any standard minimization scheme of numerical analysis

- We shall explore this approach here
  - but, there is a catch...
Basis for density-matrix energy minimization

- On the previous slide, we assumed a basis of atom-fixed atomic orbitals (AOs)
  - well suited to integration and sparse for large systems
  - unfortunately, linear equations are ill-conditioned in this nonorthogonal basis
- To obtain well-conditioned equations we orthonormalize the AO basis (Head-Gordon)
  \[ Z^T S Z = I \]
  - there are infinitely many such orthonormalizations
- The Löwdin basis retains greatest similarity with the original AOs
  \[ Z = S^{-1/2} \]
  - retains sparsity
  - traditionally evaluated by diagonalization
- We apply the iterative Newton–Schulz iterations:
  \[ Z_1 = I, \quad Z_{n+1} = \frac{1}{2} Z_n \left( 3I - Z_n^T S Z_n \right), \quad S^{-1/2} = \lim_{n \to \infty} Z_n \]
  - convergence guaranteed with scaled \( \lambda_{\text{min}} S \) that minimizes \( \| \lambda S - I \|_2 \)
  - typically 15 iterations and 50 matrix multiplications needed for convergence to \( 10^{-10} \)
Optimization in orthonormal Löwdin basis

• We begin all calculations by evaluating $S^{\pm 1/2}$ iteratively
  – example: polyalanine peptide residues (6-31G) (cost vs. number of atoms)

• Our optimization is then carried out in the Löwdin basis

$$\mathcal{E}(X) = \text{Tr} D(X)\tilde{h} + \frac{1}{2} \text{Tr} D(X)\tilde{G}(D(X))$$

by transforming all integrals in the following manner

$$\tilde{h} = S^{-1/2}hS^{-1/2}$$

• In this basis, all operations are well conditioned and no overlaps appear $\tilde{S} = I$
Newton’s method applied to the SCF energy

- We can now minimize the SCF energy in an unconstrained manner

\[ \mathcal{E}(X) = \text{Tr} D(X)\hat{h} + \frac{1}{2} \text{Tr} D(X)\hat{G}(D(X)) \]

\[ D(X) = \exp(-X)D \exp(X), \quad X^T = -X \]

- The calculation of electronic energies is a classical problem of optimization:

“Before 1940 relatively little was known about methods for numerical optimization of functions in many variables. There had been some least squares calculations carried out, and steepest descent type methods had been applied in some physics problems. The Newton method in many variables was known, and more sophisticated methods were being attempted such as the self-consistent field for variational problems in theoretical chemistry.”


- We shall begin by attempting Newton’s method
  - fast second-order local convergence
  - modified for guaranteed global convergence (trust-region method)
  - each iteration is expensive

- We shall then see how the SCF method comes back to visit us again…
Newton’s method applied to the SCF energy

• In Newton’s method, we calculate a sequence of Newton steps

\[
\text{Hessian } \times \text{ step } \rightarrow \sum_{\nu} \frac{\partial^2 E(X)}{\partial X_\mu \partial X_\nu} X_\nu = -\frac{\partial E(X)}{\partial X_\mu} \quad \leftarrow \text{negative gradient}
\]

- gradient and Hessian carry information about slope and curvature, respectively
- excellent convergence properties but each step is expensive

• Introducing projections onto the occupied and virtual orbital spaces

\[
F_n = F_{n}^{oo} + F_{n}^{ov} + F_{n}^{vo} + F_{n}^{vv}, \quad (P_o = D_n, P_v = I - D_n)
\]

the Newton matrix equation at the \( n \)th iteration become

\[
(F_{n}^{vv} - F_{n}^{oo}) X + X (F_{n}^{vv} - F_{n}^{oo}) + G^{ov}([D_n, X]) - G^{vo}([D_n, X]) = F_{n}^{vo} - F_{n}^{ov}
\]

- inexpensive
- expensive

• Each Newton matrix equation is solved iteratively in about 10 iterations
  - each Newton step requires several integral and XC reevaluations \( G([D_n, X]) \)

• Note carefully: there are both macro (outer) and micro (inner) iterations
  - each Newton iteration is a macro iteration
  - each linear Hessian transformation is a micro iteration
Some examples of Newton optimizations

- Newton density-matrix optimizations of the total electronic energy
  - water HF/DZ (red), water dimer B3LYP/DZ (yellow), adenine HF/STO-3G (blue), rhodium complex HF/STO-3G+ADZ (green)

- Convergence fast in the local region, slow but smooth in global region
  - global convergence guaranteed by the trust-region method
  - the number of Fock/KS matrix reevaluations: 24, 42, 54, 211
The structure of the Newton matrix equation

- Newton’s method converges well but it is expensive

\[
E[D(X)] = \text{Tr} D(X) \bar{h} + \frac{1}{2} \text{Tr} D(X) \tilde{G}(D(X))
\]

\[
D(X) = \exp(-X)D \exp(X), \quad X^T = -X
\]

- There are two distinct contributions to the Hessian of the Newton matrix equation

\[
(F_{nn}^{vv} - F_{nn}^{oo}) X + X (F_{nn}^{vv} - F_{nn}^{oo}) + G^{ov}([D_n, X]) - G^{vo}([D_n, X]) = F_{nn}^{vo} - F_{nn}^{ov}
\]

- the first contribution requires only matrix multiplications with \( X \)
- it arises since \( D(X) \) is nonlinear in \( X \)
- the second contribution requires Fock/KS build from \([D_n, X]\)
- it arises since \( E[D(X)] \) is nonlinear in \( D(X) \)

- Can we simplify the Newton matrix equation without destroying convergence?

- We shall explore simplifications and in particular investigate
  - omission of the expensive term
  - omission of the inexpensive term
A simplified Newton method

- Let us begin by ignoring the time-consuming part of the Newton equations:

\[(F_{vv}^n - F_{oo}^n) \mathbf{X} + \mathbf{X} (F_{vv}^n - F_{oo}^n) = F_{vo}^n - F_{ov}^n\]

- the solution of these equations is simple and fast
- about 10 iterations and about 100 matrix multiplications needed

- But energy convergence suffers since the Newton step is now approximate
  - water (red), water dimer (yellow), adenine (blue), rhodium complex (green)

- linear convergence in the local region, global convergence not guaranteed
- the number of Fock/KS matrix reevaluations: 20 (24), 50 (42), 33 (54), ∞ (211)
The Roothaan–Hall Newton method

- We have tried a simplified Newton step with mixed success:

\[(F_{nn}^{vv} - F_{nn}^{oo}) X + X (F_{nn}^{vv} - F_{nn}^{oo}) = F_{nn}^{vo} - F_{nn}^{ov}\]

- Alternative view: this the exact Newton step of an approximate energy function

\[\mathcal{E}_{RH}(X) = \text{Tr} D(X) F_n = \sum_i \epsilon_i \quad \text{← sum of orbital energies}\]

- we are in fact carrying out a Roothaan–Hall diagonalization step!

- We call this the Roothaan–Hall (RH) Newton method

- diagonalization-free
- SCF theory is simple
- good alternative to diagonalization for large systems
- 50–100 sparse matrix multiplications

![Graph showing time in RH Newton equations against the number of atoms.](image)
The DIIS step: the alternative approximate Newton step

- We obtain the RH diagonalization step by dropping the expensive part of the Hessian
  - what happens if we drop the inexpensive part instead?

\[
G^{ov}([D_n, X]) - G^{vo}([D_n, X]) = F^v_n - F^o_n
\]

- It is expensive, but we may evaluate it approximately, to good accuracy
  - during the optimization, we have collected a set of density matrices \( D_{in} = D_i - D_n \)
  - \( G(D_{in}) \) is (approximately) known in these directions

\[
G(D_{in}) = G(D_i) - G(D_n) \approx F_i - F_n = F_{in} \leftarrow \text{available}
\]

- Let us therefore project \([D_n, X]\) onto the directions \( D_{in} \)

\[
[D_n, X] \approx \sum_i c_i(X)D_{in} \implies G([D_n, X]) \approx \sum_i c_i(X)F_{in}
\]
  - this evaluation is inexpensive, requiring no new Fock/KS matrix evaluations

- The resulting step is equivalent to the DIIS step

\[
\sum_i c_i(X) (F^{ov}_{in} - F^{vo}_{in}) = F^v_n - F^o_n
\]
  - the RH/DIIS method alternates between the two approximate Newton steps
The augmented Roothaan–Hall (ARH) method

- We have seen how the full Newton step
  \[(\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) \mathbf{X} + \mathbf{X} (\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) + \mathbf{G}^{ov}([\mathbf{D}_n, \mathbf{X}]) - \mathbf{G}^{vo}([\mathbf{D}_n, \mathbf{X}]) = \mathbf{F}_{nn}^{vo} - \mathbf{F}_{nn}^{ov}\]
gives rise to the RH step and the DIIS step
  \[(\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) \mathbf{X} + \mathbf{X} (\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) = \mathbf{F}_{nn}^{vo} - \mathbf{F}_{nn}^{ov} \quad \leftarrow \text{RH diagonalization step}\]
  \[\sum_i c_i(\mathbf{X})(\mathbf{F}_{in}^{ov} - \mathbf{F}_{in}^{vo}) = \mathbf{F}_{nn}^{vo} - \mathbf{F}_{nn}^{ov} \quad \leftarrow \text{DIIS averaging step}\]

- There is no need to alternate between the two steps
  - in the two-step RH/DIIS method, we use incomplete information at each step

- Let us simplify the Newton equations by a DIIS-type projection on the expensive part
  \[(\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) \mathbf{X} + \mathbf{X} (\mathbf{F}_{nn}^{vv} - \mathbf{F}_{nn}^{oo}) + \sum_i c_i(\mathbf{X})(\mathbf{F}_{in}^{ov} - \mathbf{F}_{in}^{vo}) = \mathbf{F}_{nn}^{vo} - \mathbf{F}_{nn}^{ov} \quad \text{(JO)}\]
  - we now make full use of the available curvature information at each step
  - this is possible since we concentrate on the density matrix as the central quantity

- We call this the augmented RH method
  - a quasi-Newton method with the dominant part of the Hessian treated exactly
Examples of ARH optimizations

- The ARH method is similar to Newton globally but slower locally
  - water (red), water dimer (yellow), adenine (blue), rhodium complex (green)

- It requires the fewest number of Fock/KS evaluations (for the four systems):

<table>
<thead>
<tr>
<th>System</th>
<th>5</th>
<th>10</th>
<th>15</th>
<th>20</th>
<th>25</th>
</tr>
</thead>
<tbody>
<tr>
<td>Newton</td>
<td>24</td>
<td>42</td>
<td>54</td>
<td>211</td>
<td></td>
</tr>
<tr>
<td>RH</td>
<td>20</td>
<td>50</td>
<td>33</td>
<td></td>
<td>∞</td>
</tr>
<tr>
<td>ARH</td>
<td>9</td>
<td>15</td>
<td>15</td>
<td></td>
<td>30</td>
</tr>
</tbody>
</table>
The ARH method compared with RH/DIIS (I)

- Convergence for a variety of molecules
  - cadmium–imidazole complex (circles) B3LYP/3-21G (10/89)
  - 29-residue polyalanine peptide (crosses) B3LYP/6-31G (292/1599)
  - model B12 vitamin (triangles) HF/AhlrichsVDZ (74/428)

- Energy-error on a logarithmic scale (ARH left, RH/DIIS right)
  - RH/DIIS converges to a saddle point, ARH always to a minimum
The ARH method compared with RH/DIIS (II)

- Convergence of calculations for a variety of molecules
  - cluster of 51 water molecules (full triangles) B3LYP/cc-pVTZ (153/2958)
  - model B12 vitamin (empty circles) BP86/AhlrichsVDZ (74/428)
  - insulin (full squares) B3LYP/6-21G (786/4417)

- Energy-error on a logarithmic scale (ARH left, RH/DIIS right)
  - ARH converges smoothly, RH/DIIS oscillates (and diverges)
Review and preview

- **Large-system code conversion:**
  - **Coulomb FMM** removes the first bottleneck for large systems
    ▶ for maximum efficiency, density fitting is mandatory
    ▶ time-consuming part: exact exchange
  - **wave-function optimization** is often more difficult for large systems
    ▶ diagonalization expensive
    ▶ SCF convergence often troublesome
    ▶ the ARH method is more robust and target towards minimum
  - **response theory** poses relatively few problems
    ▶ iterative subspace algorithms used from the beginning
    ▶ convergence of the linear equations is easy (compared with SCF iterations)
    ▶ for large systems, most of the time is spent doing (nonsparse) linear algebra

- **Large-system applications:**
  - excitation energies and polarizabilities
  - indirect nuclear spin–spin coupling constants
Response theory

- We consider a system described by the one-electron density matrix

\[ D(X) = \exp(-X)D \exp(X) \quad (X = 0 \text{ for unperturbed system}) \]

and define the Hessian and metric operators in terms of their transformations

\[ E^{[2]}(X) = (F^{vv} - F^{oo})X + X(F^{vv} - F^{oo}) + G^{vo}([D, X]) - G^{ov}([D, X]) \]
\[ S^{[2]}(X) = X^{ov} - X^{vo} \]

- When perturbed by \( V_\omega \) of frequency \( \omega \), the system responds to first order as

\[ E^{[2]}(X_\omega) - \omega S^{[2]}(X_\omega) = [D, V_\omega] \quad \leftarrow \text{linear response matrix equation} \]

- perturbed density matrix: \( D_\omega = [D, X_\omega] \)
- perturbed expectation values: \( \langle \hat{A}; \hat{V}_\omega \rangle_\omega = \text{Tr} A [D, X_\omega] \)

- In the absence of a perturbation \( V_\omega = 0 \), we obtain an eigenvalue equation

\[ E^{[2]}(X_n) = \omega_n S^{[2]}(X_n) \quad \leftarrow \text{RPA matrix eigenvalue equation} \]

- transition density matrix: \( D_{0n} = [D, X_n] \)
- transition moments: \( \langle 0| \hat{A} |n \rangle = \text{Tr} A [D, X_n] \)

  - Ochsenfeld, Head-Gordon, Weber, Niklasson, and Challacombe (static properties)
Solution of the response equations

- For the solution, it is useful to consider two response equations

\[ E^{(2)}(X) - \omega S^{(2)}(X) = B \quad \leftarrow \text{full response equation} \]

\[ E_F^{(2)}(X) - \omega S^{(2)}(X) = B \quad \leftarrow \text{simplified response equation} \]

where the simplified Hessian is a good but cheap approximation to the full Hessian:

\[ E_F^{(2)}(X) = (F^{vv} - F^{oo})X + X(F^{vv} - F^{oo}) \quad \leftarrow \text{no two-electron part} \]

- its eigenvalues are the orbital energy differences \( \epsilon_a - \epsilon_i \)

- We wish solve the full response equations iteratively

\[ R_i = E^{(2)}(X_i) - \omega S^{(2)}(X_i) - B \quad \leftarrow \text{residual} \]

- new trial vectors are generated from the residual until it is sufficiently small

- For fast convergence, we precondition with the simplified response equation

\[ E_F^{(2)}(\tilde{R}_i) - \omega S^{(2)}(\tilde{R}_i) = R_i \]

- in the MO basis, \( E_F^{(2)} \) is diagonal (orbital-energy differences) and solution is trivial
- in the OAO basis, \( E_F^{(2)} \) is nondiagonal and 5–20 iterations are required for solution
Excitation energies

- The calculation of excitation energies (RPA) more difficult than that of polarizabilities
  - the eigenmatrices of the simplified Hessian provide good starting guesses:

\[ \mathbf{E}_F^{(2)}(X_{ia}) = (\epsilon_a - \epsilon_i)S^{(2)}(X_{ia}) \quad \text{where} \quad X_{ia} = C_iC_a^T, \]

\[ \begin{cases} F^{oo}C_i = \epsilon_i C_i \\ F^{vv}C_a = \epsilon_a C_a \end{cases} \]

- slow convergence of preconditioning equations (subspace problem nearly singular)

- CAMB3LYP/6-31G alanine residues
- cubic complexity with dense-matrix algebra
- linear complexity with sparse-matrix algebra
- preconditioning part dominates
- Fock/KS matrix construction dominates (not shown here)
Excitation energies: comparison with coupled-cluster theory for CO
Deficiencies of exchange–correlation functionals and potentials

- Standard pure (continuum) DFT functionals suffer from some known deficiencies
  - an incorrect asymptotic form of the exchange–correlation potential

\[
\lim_{r \to \infty} v_{xc}(r) = \lim_{r \to \infty} \frac{\delta E_{xc}[\rho]}{\delta \rho(r)} \neq -\frac{1}{r}
\]

- no integer discontinuity of the exchange–correlation potential

\[
v_{xc}^+(r) \neq v_{xc}^-(r) + \Delta_{xc}, \quad \Delta_{xc} > 0
\]

- a nonvanishing self-interaction energy

\[
J[\rho] + E_{xc}[\rho] \neq 0 \quad \text{for } 0 \leq \int \rho(r) dr \leq 1
\]

- These deficiencies give rise to problems with many properties
  - too low Rydberg excitations, charge-transfer (CT) and core excitations
  - too large polarizabilities

- These problems are alleviated by the introduction of exact exchange
  - this can be done in several different ways
Proportion of exact exchange in different functionals

![Graph showing the proportion of exact exchange in different functionals]

- HF
- B3LYP
- CAM-B3LYP
- LC

The graph illustrates the proportion of exact exchange with different functionals, showing that CAM-B3LYP and LC exhibit a higher proportion of exact exchange compared to HF and B3LYP.
Assessment of calculated excitation energies

- We have made an extensive assessment of DFT excitation energies
  - Peach et al. JCP 128, 044118 (2008)
  - 59 singlet excitations of 18 theoretically challenging main-group molecules

\[
\text{dipeptide} \quad \beta\text{-dipeptide} \quad \text{tripeptide}
\]
\[
\text{acenes (} n=1\text{–}5) \quad \text{N-phenylpyrrole} \quad \text{DMABN}\*
\]
\[
\text{PA oligomers (} n=2\text{–}5) \quad \text{N}_2 \quad \text{CO}
\]
\[
\text{H}_2\text{CO} \quad \text{HCl}
\]

*DMABS = 4-(N,N-dimethylamino) benzonitrile
## Tripeptide excitations

![Tripeptide molecule structure]

<table>
<thead>
<tr>
<th>excitation</th>
<th>type</th>
<th>PBE</th>
<th>B3LYP</th>
<th>CAM</th>
<th>exp.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$n_2 \rightarrow \pi_2^*$</td>
<td>local</td>
<td>5.58</td>
<td>5.74</td>
<td>5.92</td>
<td>5.61</td>
</tr>
<tr>
<td>$n_1 \rightarrow \pi_1^*$</td>
<td>local</td>
<td>5.36</td>
<td>5.57</td>
<td>5.72</td>
<td>5.74</td>
</tr>
<tr>
<td>$n_3 \rightarrow \pi_3^*$</td>
<td>local</td>
<td>5.74</td>
<td>5.88</td>
<td>6.00</td>
<td>5.91</td>
</tr>
<tr>
<td>$\pi_1 \rightarrow \pi_2^*$</td>
<td>CT</td>
<td>5.18</td>
<td>6.27</td>
<td>6.98</td>
<td>7.01</td>
</tr>
<tr>
<td>$\pi_2 \rightarrow \pi_3^*$</td>
<td>CT</td>
<td>5.51</td>
<td>6.60</td>
<td>7.68</td>
<td>7.39</td>
</tr>
<tr>
<td>$n_1 \rightarrow \pi_2^*$</td>
<td>CT</td>
<td>4.61</td>
<td>6.33</td>
<td>7.78</td>
<td>8.12</td>
</tr>
<tr>
<td>$n_2 \rightarrow \pi_3^*$</td>
<td>CT</td>
<td>5.16</td>
<td>6.83</td>
<td>8.25</td>
<td>8.33</td>
</tr>
<tr>
<td>$\pi_1 \rightarrow \pi_3^*$</td>
<td>CT</td>
<td>4.76</td>
<td>6.06</td>
<td>8.51</td>
<td>8.74</td>
</tr>
<tr>
<td>$n_1 \rightarrow \pi_3^*$</td>
<td>CT</td>
<td>4.26</td>
<td>6.12</td>
<td>8.67</td>
<td>9.30</td>
</tr>
</tbody>
</table>
### Statistics for 59 excitation energies

<table>
<thead>
<tr>
<th></th>
<th>PBE</th>
<th>B3LYP</th>
<th>CAM</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>local</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\overline{\Delta}$</td>
<td>-0.31</td>
<td>-0.15</td>
<td>0.02</td>
</tr>
<tr>
<td>$</td>
<td>\overline{\Delta}</td>
<td>$</td>
<td>0.33</td>
</tr>
<tr>
<td>$\overline{\Delta}_{std}$</td>
<td>0.27</td>
<td>0.26</td>
<td>0.27</td>
</tr>
<tr>
<td>max $</td>
<td>\Delta</td>
<td>$</td>
<td>0.83</td>
</tr>
<tr>
<td><strong>Rydberg</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\overline{\Delta}$</td>
<td>-1.84</td>
<td>-1.11</td>
<td>-0.50</td>
</tr>
<tr>
<td>$</td>
<td>\overline{\Delta}</td>
<td>$</td>
<td>1.84</td>
</tr>
<tr>
<td>$\overline{\Delta}_{std}$</td>
<td>0.30</td>
<td>0.23</td>
<td>0.18</td>
</tr>
<tr>
<td>max $</td>
<td>\Delta</td>
<td>$</td>
<td>2.24</td>
</tr>
<tr>
<td><strong>CT</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\overline{\Delta}$</td>
<td>-2.60</td>
<td>-1.35</td>
<td>-0.18</td>
</tr>
<tr>
<td>$</td>
<td>\overline{\Delta}</td>
<td>$</td>
<td>2.60</td>
</tr>
<tr>
<td>$\overline{\Delta}_{std}$</td>
<td>1.37</td>
<td>0.86</td>
<td>0.31</td>
</tr>
<tr>
<td>max $</td>
<td>\Delta</td>
<td>$</td>
<td>5.04</td>
</tr>
</tbody>
</table>
A diagnostic for excitation energies

- At present, DFT functionals for excitation energies are unsatisfactory
  - valence excitations are fine (off by a few tenths of an eV)
  - Rydberg, CT and core excitations are not poor
- When do we need to exercise caution? When can we trust our results?
  - problems arise when excitations occur between orbitals with no spatial overlap
- Excitations represented by rotations $\kappa_{ai}$ between all occupied and virtual orbitals
  - to measure spatial overlap between $s$ and $i$, we introduce
    \[
    O_{ia} = \int |\phi_i(r)| |\phi_a(r)| \, dr
    \]
    - without moduli all overlaps would be trivially zero
- We now measure spatial overall using the quantity
  \[
  \Lambda = \frac{\sum_{ai} \kappa_{ai}^2 O_{ai}}{\sum_{ai} \kappa_{ai}^2}, \quad 0 \leq \Lambda \leq 1
  \]
  - a small value of $\Lambda$ signifies a long-range excitation
  - a large value of $\Lambda$ signifies a short-range excitation
- Peach et al. JCP 128, 044118 (2008)
Excitation energies against $\Lambda$ for PBE

- With this GGA functional, performance for $\Lambda < 0.6$ is erratic
  - local excitations ($\Lambda > 0.4$) are well reproduced, within a few tenths of an eV
  - Rydberg excitations ($\Lambda < 0.3$) are systematically underestimated by a few eVs
  - CT excitations cover a surprisingly large $\Lambda$ range (DMABN)
With the introduction of exact exchange, performance improves

- **local excitations** are lifted and cover a slightly larger $\Lambda$ range
- errors in **Rydberg excitations** are reduced by a factor of two
- likewise, an improvement in **CT excitations** is observed
Excitation energies against $\Lambda$ for CAM-B3LYP

- Coulomb attenuation gives a uniform description of local and nonlocal excitations!
  - however, errors as large as 1 eV are still observed
In general, BLYP performs in much the same manner as LDA for excitation energies:
- the BLYP excitation energies are typically within 1 or 2 m\(E_h\) of the LDA energies
- the differences between LDA and BLYP are smaller for longer chains
- BLYP increases the triplet–singlet separation slightly

Plots of the LDA/6-31G and BLYP/6-31G excitation energies in the polyynes:

In general, LDA and GGA functionals underestimate excitation energies
- for HF, CO, and H\(_2\)O, the underestimation is 15% to 20%
- this problem is exacerbated in extended conjugated systems
- the introduction of nonlocality (beyond GGA) may be very important in providing quantitatively correct excitation energies, in particular in conjugated systems
CAMB3LYP/6-31G excitation energies of alanine residue peptides

- lowest excitation energy almost identical with lowest Hessian eigenvalue
- CAMB3LYP has 19% short-range and 65% long-range exact exchange
B3LYP/6-31G excitation energies of alanine residue peptides

- B3LYP (with 20% exact exchange) behaves differently, in an unphysical manner
  - beyond 27 alanine residues, the lowest excitation energy becomes very small
  - RH/DIIS converges to an excited state (first-order saddle point)

- for LDA and GGA functionals, the crossing occurs even earlier
  - local DFT functionals unable to account for long-range exchange in large molecules
The B3LYP energy minimum behaves in a strange manner
- the lowest Hessian eigenvalue goes through a minimum for 27 residues
- the RH/DIIS solution behaves smoothly with increasing chain length
B3LYP/6-31G polyalanine polarizabilities

Polyalanine, dipole polarizability

Energy / a.u.

Alanine residues
To illustrate, we have calculated longitudinal polarizabilities in linear polymeric chains
- HF and DFT $\alpha$ and $\alpha/N$ in 6-31G basis, plotted against the number of carbons $N$

- The alkenes are about an order of magnitude more polarizable than the alkanes
  - all models agree on alkanes ($\alpha/N$-limit: HF 14.4; LDA 16.3)
  - widely different results for alkenes ($\alpha/N$-limit: HF 97; LDA 427)
The importance of exact exchange for longitudinal polarizabilities

- Without a good description of long-range exchange, the systems become too polarizable

- the Hartree model neglects all exchange and overestimates by a factor of eight
- pure DFT has a poor long-range exchange and overestimates by a factor of four
- hybrid functionals improve the situation, introducing some exact exchange
- compromise solution: standard DFT at short range, full exchange at long range
Asymptotic behaviour of group polarizabilities

• How does the group polarizability converge towards the infinite limit?
  \[
  \bar{\alpha}_\infty - \bar{\alpha}_N = e N^{-1} + \mathcal{O}(N^{-2}) \quad \text{Kudin et al., JCP 122, 134907 (2005)}
  \]
  – this behaviour is universal, holding at all levels of theory

• Log–log plots of \(\bar{\alpha}_\infty - \bar{\alpha}_N\) for alkanes and alkenes:

  – limit obtained by extrapolation \(\bar{\alpha}_\infty = (\alpha_N - \alpha_M)/(N - M)\)
  – straight lines of slope \(-1\) superimposed through the points at \(N = 350\)

• The asymptotic region is reached with \(C_{30}H_{62}\) (alkanes) and \(C_{60}H_{62}\) (alkenes)
  – alkane \(\bar{\alpha}_\infty\) predicted to within 1% from \(C_{30}H_{62}\)
  – alkene \(\bar{\alpha}_\infty\) predicted to within 1% from \(C_{60}H_{62}\) for HF and from \(C_{150}H_{152}\) for LDA
Polarizabilities and hyperpolarizabilities in BN sheets

Response Property as a function of BN units

Hyperpolarizability (au.) vs. Units

Linear Response

Quadratic Response

Polarizability (au.)
Indirect nuclear spin–spin coupling constants

- With each nucleus in a molecule, there is an associated magnetic moment $M_P$:
  - their direct interactions vanish in isotropic media
  - the residual indirect interaction arises from hyperfine interactions with the electrons $\approx 10^{-16} E_h \approx 1 \text{ Hz}$

- DFT has created revolution in the calculation of spin–spin coupling constants
  - simulated 200 MHz NMR spectra of vinyllithium
DFT can be applied to large molecular systems such as valinomycin (168 atoms)

- there are a total of 7587 spin–spin couplings to the carbon atoms in valinomycin
- below, we have plotted the magnitude of the reduced LDA/6-31G coupling constants on a logarithmic scale, as a function of the internuclear distance:

- the coupling constants decay in characteristic fashion, which we shall examine
- most of the indirect couplings beyond 500 pm are small and cannot be detected
Summary

• We have considered some computational tasks in SCF theory
  – orthogonalization of the atomic orbitals
  – minimization of the SCF energy in an orthonormal basis

• Several approaches may be taken to the SCF energy minimization
  – full second-order Newton optimization
  – Roothaan–Hall (RH) Newton minimization (equivalent to diagonalization)
  – augmented RH (ARH) (reuses old Fock matrices)

• The ARH method is superior to RH/DIIS
  – requires one Fock/KS-matrix evaluation at each iteration
  – reuses efficiently information present in old Fock/KS matrices
  – converges by design to the ground state, smoothly and robustly

• For sufficiently sparse AO matrices, all tasks can be carried out in linear time
  – the Fock/KS matrix is reevaluated no more often than in MO theory
  – the remaining work is rich in matrix multiplication
  – well suited to parallelization