Modern Electronic Structure Methods: 1. Density functional theory

MARTIN HEAD-GORDON,

Department of Chemistry, University of California, and Chemical Sciences Division, Lawrence Berkeley National Laboratory Berkeley CA 94720, USA

Outline

- 1. Overview of density functional theory
- 2. Fast methods for Fock matrix formation
- 3. The diagonalization problem
- 4. Excited states and time-dependent DFT



Chemistry and mathematics

"Every attempt to employ mathematical methods in the study of chemical questions must be considered profoundly irrational and contrary to the spirit of chemistry. If mathematical analysis should ever hold a prominent place in chemistry—an aberration which is happily almost impossible– it would occasion a rapid and widespread degeneration of that science."

Auguste Comte, 1830

Quantum mechanics and chemistry

"...in the Schrödinger equation we very nearly have the mathematical foundation for the solution of the whole problem of atomic and molecular structure"

but...

"... the problem of the many bodies contained in the atom and the molecule cannot be completely solved without a great further development in mathematical technique."

G. N. Lewis, J. Chem. Phys. 1, 17 (1933)



Complexity

 Many-body Schrödinger equation is a partial differential equation in 3n unknowns- the positions of the electrons.
 Exact (brute force) solution will scale approximately exponentially with the number of electrons.

- When done in a given basis of 1-electron expansion functions, this is "full configuration interaction" (FCI).
 Largest FCI's involve many-body expansions containing billions of terms, for molecules with 2 or 3 of atoms (!).
- · Approximations are imperative. Accuracy vs feasibility.

Fundamental approximations

• (1) The one-body problem ("atomic orbital basis")

- One-electron functions are expanded in a finite basis.
 We use atom-centered functions with some maximum
- angular momentum (L) on atoms of a given period. • Standardized basis sets of increasing L are available.
- (2) The n-body problem ("electron correlation method")
 More about this in a minute mean field is the simplest
- possione nearment.
- Well-defined electronic structure models are completely specified by these two approximations.

Branches of the family tree

• Wavefunction-based electronic structure theory:

- · Minimize the energy by varying the wavefunctio
- Tremendously complicated unknown function
 - $\Psi = \Psi(\underline{r}_1, \underline{r}_2, \dots, \underline{r}_n)$
- Modeling the wavefunction yields "model chemistries"

• Density functional theory

- The unknown is very simple: $\rho = \rho(\tau)$
- Hohenberg-Kohn theorem guarantees that: $E = E \{ \rho(\tau) \}$
- True functional is unknown and probably unknowable
- Modeling the functional gives DFT model chemistries.

A brief overview of density functional theory

- First Hohenberg-Kohn theorem (1965):
- Proof by contradiction: let H_1 and H_2 have the same $\rho(\mathbf{r})$
 - Use Ψ_2 as trial function in H_1 problem
 - Use Ψ_1 as trial function in H_2 problem

$$\begin{split} E_{i}\left[\Psi_{2}\right] &= \left\langle \Psi_{2}\left|\hat{T}+\hat{\mathcal{V}}^{cc}\right|\Psi_{2}\right\rangle + \int v_{i}\rho dr > \left\langle \Psi_{1}\left|\hat{T}+\hat{\mathcal{V}}^{cc}\right|\Psi_{1}\right\rangle + \int v_{i}\rho dr \\ E_{i}\left\{\Psi_{1}\right\} &= \left\langle \Psi_{1}\left|\hat{T}+\hat{\mathcal{V}}^{cc}\right|\Psi_{1}\right\rangle + \int v_{i}\rho dr > \left\langle \Psi_{2}\left|\hat{T}+\hat{\mathcal{V}}^{cc}\right|\Psi_{2}\right\rangle + \int v_{i}\rho dr \end{split}$$

Contradiction:

A brief overview of density functional theory

- First Hohenberg-Kohn theorem (1965):
 - 1:1 mapping between ground state electron densities and Hamiltonian
 - Ground state energy E is determined directly from the Hamiltonian
 - Hence *E* is given in terms of the density, $\rho(\mathbf{r})$.
- A formal construction exists for the exact functional, $E = E \{ p(\underline{r}) \}$ • Constrained search over all wavefunctions yielding $p(\mathbf{r})$ (!!!)
- So, in practice the functional must be modeled.
- Given a functional, and an external potential (nuclear field) ρ(r) is found by minimizing over allowed densities.

Construction of model density functionals

- Need to model kinetic, exchange and correlation function
- Largest energy contribution is the kinetic energy.
 No satisfactory kinetic energy functional vet exists
- Kohn-Sham framework (a beautiful sidestep):
 - Use the kinetic energy of a non-interacting system with the same electron density (a Hartree-Fock type wavefunction).

 - Kohn-Sham computational cost: similar to mean-field Hartree-Fock
 - Still cheap enough to apply to large systems.

Modern Kohn-Sham density functionals

- Local density approximation (LDA): 1960's, 1970's
 - Functional depends only on the density at each point, *ρ*(**r**)
 LDA overbinds as much as Hartree-Fock (mean field) underbinds!
- Generalized gradient approximations (GGA's): 1988
 - Functional depends on density $\rho(\mathbf{r})$ and its gradients $\nabla \rho(\mathbf{r})$ at each \mathbf{r}
 - Greatly improved results! 4-6 kcal/morentor for BLYP, PW91, PBE
- Exact exchange mixing (adiabatic connection): 1992
 - Mix some Hartree-Fock exchange with GGA's (Becke)Best yet! 2-3 kcal/mol error for B3LYP

Computational steps in DFT calculations

- Given a guess at the density matrix, P, (describing the arrangement of the electrons), make the Hamiltonian matrix describing their effective interactions, F=F(P).
 - Electron-electron, electron-nuclear and kinetic energy terms
 - Electron-electron terms are computationally dominant: A²
 - Includes Coulomb, (sometimes) exact exchange, and exchange-correlation (XC) terms.
- (2) Diagonalize the effective Hamiltonian to obtain the eigenvalues and eigenvectors, and use them to make P.
- (3) If the new **P** is different to the old **P**, go back to (1).

Convergence with atomic orbital basis set

- Depends on the highest angular momentum (L).
- Hartree-Fock (& DFT)
 - Roughly exponential convergence with L... rapid.
 - Still need at least f functions (on C,N,O...) for B3LYP thermochemist

Wavefunction-based correlation methods.

- L^o convergence... very slow.
- Cost increases approximatery as 2 ... very fas
- Achieving chemical accuracy feasibly is challenging.
 Particle number and angular momentum bottlenecks

Summary of present status of DFT

- The only method in large-scale use for electronic structure studies in condensed matter (using plane-wave expansions)
- The predominant method in use for electronic structure studies of molecules, surfaces, and nanomaterials

 Basis set expansions are usually Gaussian "atomic-orbital-like" functions
 Standardized Gaussian basis sets are available for virtually all elements.
- Strengths of DFT:
 It is the sweet spot between accuracy and feasibility (already)
- Weaknesses of DFT:
 Not systematically improvable
 Computational cost is still high relative to empirical methods











Prospects for significantly better functionals?

- How to make progress without destroying the simplicity of present-day density functional methods?
 - The functionals are (mostly) local in space
- Find local quantities which in some sense are diagnostics for cases where delocalized exchange holes exist
 - Quantities which depend on the kinetic energy density, and/or density hessians
 To date, the improvements over B3LYP are not sufficient to be useful.
- Eliminate self-interaction using the optimized effective potential (OEP).
- Exchange potential is now that which is due to the Kohn-Sham orbitals
 Requires the development of new correlation functionals, which are non-local.

Outline

1. Overview of density functional theory

2. Fast methods for Fock matrix formation

- Gaussian basis sets and matrix elements
- Coulomb interactions -- short and long-rang
- A mention of exchange and correlation
- 3. The diagonalization problem
- 4. Excited states and time-dependent DFT



Linear scaling evaluation of Coulomb interactions

- · We divide Coulomb energies and gradients into:
 - (1) short-range (near-field) terms from overlapping functions(2) long-range (far-field) terms from nonoverlapping functions
- For the far-field, we employ the Continuous Fast Multipole Method (CFMM):

 - And Crivity Context vizes on grange interactions into interactions interactions into interactions interaction
- For the near-field, we use an analytical J-matrix engine
 - Energies: Y. Shao, MHG, Chem. Phys. Lett. 323, 425 (2000)
 Forces: Y. Shao, MHG, C.A. White, J. Chem. Phys. 114, 6572 (2001)

One-particle basis sets in quantum chemistry

- Gaussian atomic orbitals are most commonly used: Analytical matrix element evaluation is efficie

$$|a_x, a_y, a_z; \alpha| = N (\mathbf{x} - A_x)^{\lambda} (\mathbf{y} - A_y) (\mathbf{z} - A_z)^{\lambda} e^{-\alpha (\mathbf{x} - \mathbf{y})}$$

$$N_x = \left(\frac{2\alpha}{\lambda}\right)^{\lambda} (4\alpha)^{\frac{(\alpha_x + \alpha_y + \alpha_z)}{2}} (2a_x - 1)!! (2a_x - 1)!! (2a_x - 1)!!$$

 $(a_z) = \sum_{i=1}^{n} c_i \left[a_x, a_y, a_z; \alpha_i \right]$

Contracted Gaussians:
$$a_x, a_y$$

- Basis sets are standardized. The 3 lowest levels:
 Minimal basis set (STO-3G) (5 functions per C)
 Split valence basis (3-21G) (9 functions per C)
 Polarized split valence basis (15 per C)

Shells and shell-pairs

- For efficiency, basis sets are composed of shells
 - A shell is a set of basis functions having common angular momentum (L), exponents and contraction coefficients
- Shell pairs: are products of separate shells



- If there are O(N) functions in the shell list, then, naively,
- The product of Gaussian functions is a Gaussian at P
 - $e^{-\alpha(\mathbf{r}-\mathbf{A})^2}e^{-\beta(\mathbf{r}-\mathbf{B})^2} = e^{-\frac{\alpha\beta}{(\alpha+\beta)}(\mathbf{A}-\mathbf{B})^2}e^{-(\alpha+\beta)(\mathbf{r}-\mathbf{P})^2}$

$$\mathbf{P} = \frac{\alpha \mathbf{A} + \beta \mathbf{B}}{\alpha + \beta}$$

- Pre-factor dies off with separation of the product functions, so:
- - - There are O(N²) two-electron integrals (the Coulomb problem)

Two-electron integrals in DFT calculations

- The effective Hamiltonian (Fock operator) is:
 - $F_{\mu\nu} = H_{\mu\nu} + J_{\mu\nu} + K_{\mu\nu} + (XC)_{\mu\nu}$

 - Exchange-correlation (**XC**) is a 3-d numerical quadrature, which can be evaluated in linear scaling effort.
- Two-electron integrals arise in the Coulomb (J) matrix:
 - $J_{\mu\nu} = \sum \langle \mu \nu | \lambda \sigma \rangle P_{\lambda\sigma}$

 $\langle \mu v | \lambda \sigma \rangle = \iint d\mathbf{r}_1 d\mathbf{r}_2 \omega_{\mu} (\mathbf{r}_1) \omega_{\nu} (\mathbf{r}_1) |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \omega_{\lambda} (\mathbf{r}_2) \omega_{\sigma} (\mathbf{r}_2)$

• And in the exact exchange (K) matrix:

 $K_{\mu\nu} = \sum \langle \mu \lambda | \nu \sigma \rangle P_{\lambda\sigma}$

Short-ranged density matrices

- The K-matrix is needed in hybrid DFT methods such as B3LYP.
- K is short-range, if the density matrix, ρ , is also short-range:

$$K_{\mu\nu} = \sum_{i} \langle \mu \lambda | \nu \sigma \rangle P_{i}$$

$\langle \mu v | \lambda \sigma \rangle = \iint d\mathbf{r}_1 d\mathbf{r}_2 \omega_\mu (\mathbf{r}_1) \omega_\nu (\mathbf{r}_1) |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \omega_\lambda (\mathbf{r}_2) \omega_\sigma (\mathbf{r}_2)$

- A non-zero contribution to K is only obtained if:

Fast multipole method (FMM)

- An O(N) method for summing O(N²) r⁻¹ interactions

 - L.Greengard, V.Rokhlin, J. Comput. Phys. 60, 187 (1985) L.Greengard, "The Rapid Evaluation of Potential Fields in Particle Systems" (MIT Press, 1987).
 - Innumerable subsequent contributions by many groups. I shall follow our own presentation here:

 - C.A.White, MHG, J. Chem. Phys. 105, 5061 (1997) C.A.White, MHG, Chem. Phys. Lett. 257, 647 (1996)

FMM fundamentals

- Based on the multipole expansion of r⁻¹:
- $\frac{1}{|\mathbf{r}-\mathbf{a}|} = \sum_{l=0}^{\infty} P_l(\cos\gamma) \frac{a^l}{r^{l+1}} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \frac{(l-|m|)!}{(l+|m|)!} \frac{a^l}{r^{l+1}} P_{lm}(\cos\alpha) P_{lm}(\cos\theta) e^{-im(\beta-\phi)}$ • Define chargeless multipole-like and Taylor-like moments:
 - $O_{lm}(\mathbf{a}) = \frac{1}{(l+|m|)!} a^l P_{lm}(\cos\alpha) \cdot e^{-im\beta}$

$$M_{lm}(\mathbf{r}) = \left(l - |m|\right)! \frac{1}{r^{l+1}} P_{lm}(\cos\theta) e^{i\theta}$$

· The expansion of an inverse distance is now simply:

$$\frac{1}{|\mathbf{r}-\mathbf{a}|} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} O_{lm}(\mathbf{a}) M_{lm}(\mathbf{r})$$

Objective of the FMM

- The FMM is based on collectivizing interactions by translating expansions to common origins and summing.
- It is a framework in which the collectivization is done automatically in linear scaling work with bounded error.
- This is accomplished by:
 - Developing operators to translate and interconvert multipole and Taylor expansion coefficients

 - Applying these operators to a division of space into boxes as a binary tree structure.



- Operator A translates the origin of a multipole expansion
 - $O_{lm}(\mathbf{a} + \mathbf{b}) = \sum_{k=1}^{k} \sum_{ik=1}^{k} A_{ik}^{lm}(\mathbf{b}) O_{ik}(\mathbf{a})$ $A_{jk}^{lm}(\mathbf{b}) = O_{l-j,m-k}(\mathbf{b})$
- · Operator B converts a multipole expansion about a local center into a Taylor expansion about a distant center
 - $M_{lm}(\mathbf{a}-\mathbf{b}) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} B_{jk}^{lm}(\mathbf{b}) O_{jk}(\mathbf{a})$ $B_{ik}^{lm}(\mathbf{b}) = M_{i+l,k+m}(\mathbf{b})$
- Operator C translates the origin of a Taylor expansion $M_{lm}(\mathbf{r}-\mathbf{b}) = \sum_{k=1}^{m} \sum_{jk=1}^{m} C_{jk}^{lm}(\mathbf{b}) \cdot M_{jk}(\mathbf{r}) \qquad C_{jk}^{lm}(\mathbf{b}) = A_{lm}^{jk}(\mathbf{b}) = O_{j-l,k-m}(\mathbf{b})$
- The scaling with angular momentum is clearly O(L⁴)

FMM step A: form and translate multipoles Pass 1 · Divide space in a binary tree





FMM final step D: assemble potential

• "Far-field" contributions

• "Near-field" contributions

- Doubling the number of particles hence adds 1 tier

Thinking about linear scaling in the FMM

- of particles (in 1-d)
- · To keep the number of particles per lowest level box constant, we
- For steps A, B, C of the FMM, total work scales with the number of boxes for the translations
- In step D, the work scales with the number of non-well-separated charges (per charge) multiplied by the number of charges

Accelerating the translation operators

- The L⁴ scaling of the translations can be reduced to L³ by choosing special axes such that translation is along the quantization axis, z (ie. $\theta=0$; $\phi=0$).
- The translation operators then simplify to:

•
$$\delta$$
 functions means $O(L^3)$ $A_{jk}^{in}(\mathbf{b}) = \frac{1}{(l-j+|m-k|)!} b^{l-j} \delta_{m-k,0}$

 $B_{jk}^{lm}(\mathbf{b}) = (j+l-|k+m|)! \frac{1}{b^{j+l+1}} \delta_{k+m,0}$

 $C_{jk}^{lm}(\mathbf{b}) = \frac{1}{(j-l+|k-m|)!} b^{j-l} \delta_{k-m,0}$

- Modified translations:

 - Rotate back to original axes

Continuous fast multipole method (CFMM)

- Generalize the FMM to charge distributions with extent. C.A.White et al, Chem. Phys. Lett. 230, 8 (1994)
 C.A.White et al, Chem. Phys. Lett. 253, 268 (1996)
 See also CFMM-related work by Scuseria et al (1996--).
- Many other efficient alternatives exist
- Tree code work by Challacombe (1996--).Other multipole-based methods (Yang, Friesner, etc)
- distributions separated by the sum of their extents behave as nonoverlapping to target precision.

Well-separatedness and extent

- The extent of a distribution affects what other distributions it may interact with via multipoles.
- Before, well-separatedness (WS) was a global parameter
- Now it cannot be because charge distributions may have greatly varying extents, r_{ext} . We modify the definition of well-separateness to apply
- to each charge distribution. For box length *l*:

WS = max
$$\left(2\left\lceil\frac{r_{ext}}{l}\right\rceil, WS_{ref}\right)$$

Extent of Gaussian charge distributions

• The Coulombic interaction of two spherical Gaussian charge distributions can be represented in closed form as

$$V = \frac{1}{|\mathbf{R}|} \cdot erf\left(\sqrt{\frac{pq}{p+q}} |\mathbf{R}|\right)$$

- The *erf* factor rapidly approaches 1 with increasing r, and the two distributions then interact as point charges.
- The CFMM achieves linear scaling by finding the point at which two distributions interact as point multipoles.

$$r_{ext} = \frac{1}{2} \sqrt{\frac{2}{p} \ln(\varepsilon)} \qquad \qquad r_{ext} = \frac{1}{2} \sqrt{\frac{2}{p} \ln(\varepsilon) - AB^2}$$

Absolute (right) rather than relative (left) precision is OK

Charge distributions in DFT calculations

- Two-electron interactions are between the shell pair list
- This generates a very large number of charges, very roughly on the order of 100 times the number of basis functions, which itself may be on the order of hundreds to thousands.
- The significant shell pair list must have their extents determined.
- We now need to sort these shell pairs by their position and also by their extent. The extent will determine what other distributions they can interact with.
- This requirement adds another dimension to the tree.



Operation of the generalized tree structure

- Branch each level of the tree according to WS values
- Step A: form and translate multipoles
 - Place distributions according to position and WS value in the lowest level of the tree, and make multipoles
 - Pass distributions up the free, naiving the w.S. value at each the
- Step B: external to local translation
 - Perform external to local translation on each level of the tree.
 Determine well-separateness as the average of the WS values of the branches to decide whether or not to translate.
- Step C: translate Taylor expansions
 Pass Taylor coefficients down the tree to every h

Final step (D) of the CFMM

• Evaluation of "far-field" contributions

- Use the Taylor moments in the lowest level box with the appropriate WS value...
- Evaluation of the "near-field" contributions
 - We now have a definition of the near-field interactions that must be evaluated explicitly.
 - This can be done via conventional two-electron integral evaluation, or by specialized methods that are more efficien
 - We now turn to a discussion of this problem before showin timings for representative systems.

Applying CFMM to the (far field) Coulomb force

- Consider displacement of an atomic center A:
 - By definition, the far-field contribution is unaltered...
 Good news! Local Taylor expansions remain the same (ie. Steps A,B,C are unaltered)
- Steps A,B,C are unaltered)Changes in the finest level multipole moments (to be
- contracted with the unaltered Taylor expansions) occur in 3 ways due to displacement of A:
 - The shell pair pre-factor changes
 - Pre-processed density matrix elements change (because they depend on the AB distance)
 - The effective center P itself change
- But this only affects the last step (D), and is inexpensive





Adding angular momentum by recursion

• To permit re-use of intermediates, recurrence-based formulations are natural. Our previous example can be re-cast in this way as:

$$\begin{bmatrix} a+1, s \mid ss \end{bmatrix}^{(m)} = \left(\frac{\beta}{\alpha+\beta}\right) (\mathbf{A}-\mathbf{B})_{i} \begin{bmatrix} as \mid ss \end{bmatrix}^{(m)} + \left(\frac{1}{\alpha+\beta}\right) R_{i} \begin{bmatrix} as \mid ss \end{bmatrix}^{(m+1)}$$
$$-a_{i} \left\{ \left(\frac{1}{\alpha+\beta}\right) \begin{bmatrix} as \mid ss \end{bmatrix}^{(m)} - \left(\frac{1}{\alpha+\beta}\right)^{2} \begin{bmatrix} as \mid ss \end{bmatrix}^{(m+1)} \right\}$$

 Obviously the recurrences become more complicated when higher angular momentum is involved... a variety of efficient schemes exist (McMurchie-Davidson, Obara-Saika and offshoots, Prism). Explicit evaluation of J-matrix contributions

• The Coulomb contribution to the Fock matrix: $J_{\mu\nu} = \sum \langle \mu\nu | \lambda\sigma \rangle P_{\lambda\sigma}$

$$\langle \mu v | \lambda \sigma \rangle = \iint d\mathbf{r}_1 d\mathbf{r}_2 \omega_\mu (\mathbf{r}_1) \omega_v (\mathbf{r}_1) |\mathbf{r}_1 - \mathbf{r}_2|^{-1} \omega_\lambda (\mathbf{r}_2) \omega_\sigma (\mathbf{r}_2)$$

- Consider contributions from individual shell quartets
 E.g. 1296 dddd integrals for a dddd shell quartet contribute to 36 elements of the J matrix (a single dd shell pair) (Cartesian d functions)
- · Objective: minimize the cost of this computation.

J matrix engine approach

- The 2-electron integrals are (bulky) $O(L^8)$ intermediates leading to a (small) set of $O(L^4)$ J-matrix contributions.
- Is it possible to find a more suitable set of intermediates for the specific purpose of producing the J-matrix?
 A different path through the recurrence relations?
 Reduce intermediates by early contraction with density
 - Move work from shell quartet to shell pair loops
 - C.A.White and MHG, J. Chem. Phys. 104, 2620(1996).Y.Shao and MHG, Chem. Phys. Lett. 232, 425 (2000).

A McMurchie-Davidson J Engine (Yihan Shao)

- Consider a primitive shell quartet [ab|cd]:
 P (Gaussian product center of bra functions A and B), and Q (center of ket functions C and D) are its natural centers.
- McMurchie-Davidson recurrence relations permit efficient build-up of angular momentum at P and Q.
- Transfer relations to move angular momentum to/from center P and centers A and B are independent of the ket.
 Work associated with these steps can be removed to shellpair loops (Ahmadi and Almlof, 1995)

A McMurchie-Davidson J Engine (Yihan Shao)

- Pre-processing (shell pair loops: free) • Preprocess density matrix to center Q from C and D
- Shell quartet loop calculations of near-field J-matrix:
 - Build the fundamental [0]^(m) integrals.
 Create angular momentum at Q only.

 - Contract with the preprocessed density matrix at Q
- Post-processing (shell pair loops: free)

Floating point operation counts (Yihan Shao) (expressed per J matrix contribution)

· ·		
	present	HGP
(pp pp)	28.1	160.2
(dd dd)	69.3	792.3
(ff ff)	141.4	2630.3
(gg gg)	226.2	

HGP is J-matrix evaluation from the two-electron integrals.

Generalization to the Coulomb force (Yihan Shao) • The algorithm can be generalized for the Coulomb force.

· Floating point operation counts for several shell quartets:

	present	HGP	ratio
(pp pp)	509	8,030	15.7
(dd dd)	4,105	154,030	37.5
(ff ff)	18,504	1,284,000	69.4







- Exchange cost, even for good insulators, is higher than the Coulomb cost (because the J-engine cannot be used for exchange).

Linear scaling quadrature for exchange-correlation

- Exchange-correlation (XC) functionals are complicated functions that depend on: ρ_α, ρ_β, ∇ρ_α, ∇ρ_α, ∇ρ_α, ∇ρ_β, ∇ρ_β, ∇ρ_β
- Matrix elements of these XC functionals cannot be analytically evaluated (by contrast with ordinary operators)
- We use the "standard grids" (SG0 and SG1) of Gill et al, which are atom-centered grids (~10 3 -10 4 points per atom).
- Locality of the basis functions permits linear scaling evaluation of the quadrature for these matrix elements.

 Substantial improvements in Q-Chem due to Shawn Brown, Jing Kong, using an "incremental" approach...

Density functional methods: typical timings

- BLYP/6-31G** (25 functions per wate
- Timings in minutes on a 500MHz Alpha (you'll do better!)

	H-build	diagonalize	force
(H ₂ O) ₄₀ (N=1000)	9.6	0.5	17.9
(H ₂ O) ₈₀ (N=2000)	25.9	8.5	47.0
(H ₂ O) ₁₂₀ (N=3000)	47.3	39.5	78.8
(H ₂ O) ₁₆₀ (N=4000)	66.8	89.6	115.8

Calculations performed by Yihan Shao (Berkeley)

Outline

- 1. Overview of density functional theory
- 2. Fast methods for Fock matrix formation

3. The diagonalization problem

- Short-ranged nature of the density matrix
- Curvy steps and Chebyshev polynomials
- 4. Excited states and time-dependent DFT

Solving for the energy and density matrix

• The objective is to minimize the energy. The tight-binding problem represents one SCF iteration:

$E = Tr(\hat{\rho}\hat{f})$

- The density cannot be minimized without constraints. There are two types:
 - <u>Idempotency</u> (purity): ensures that the density matrix derives from a single Kohn-Sham determinant

$\hat{\rho}\hat{\rho} = \hat{\rho}$

 $Tr(\hat{\rho}) = n$

Density matrix locality in real space (Roi Baer) • R. Baer, MHG, JCP 107, 10003 (1997) • For ordered systems, density matrix elements decay exponentially with distance. • The range (for 10^{-D} precision) is bounded by: • $W(P) = D \sqrt{\frac{3k^2}{4m_e(\Delta L)_{row}}}$ • Decay length proportional to inverse square root of gap. • Electronic structure of good insulators is most local ⇔ occupied orbitals can be well localized.

- Metals and small gap semiconductors are more nonloca
 - ↔ occupied orbitals *cannot* be well localized



Computational implications

- (1) Only a linear number of density elements are significant on length scales longer than the decay length.
- (2) When the linear scaling regime is reached depends on the *effective* dimensionality, d, going roughly as n^d

10 significant neighbors \Rightarrow 10, 100, 1000 atoms in 1,2,3-d

- (3) An ansatz will be required to develop effective fast methods on systems below the linear scaling regime.
- (4) We must work directly in a localized basis to reflect real space locality in sparse matrices.

General approaches to linear scaling

A good review: S. Goedecker, Rev. Mod. Phys. 71, 1085 (1999)

How do we avoid N^3 diagonalization & get linear scaling?

Solve directly for the density matrix
 LNV canonical purification Chebyshev

(2) Solve for localized orbitals via domain condition

(3) Divide and conquer by splitting the system

We shall discuss our most recent work on the first approach...

If we are given a valid guess density matrix... P.P. = P. m(P) = n

- Then unitary transforms preserve idempotency and electron count $\mathbf{U}^{\dagger}\mathbf{U}=\mathbf{1} \qquad \mathbf{P}=\mathbf{U}^{\dagger}\mathbf{P}_{s}\mathbf{U}$

 $\mathbf{PP} = (\mathbf{U}^{\dagger}\mathbf{P},\mathbf{U})(\mathbf{U}^{\dagger}\mathbf{P},\mathbf{U}) = \mathbf{U}^{\dagger}\mathbf{P},\mathbf{P},\mathbf{U} = \mathbf{U}^{\dagger}\mathbf{P},\mathbf{U} =$

 $\operatorname{tr}(\mathbf{P}) = \operatorname{tr}(\mathbf{U}^{\dagger}\mathbf{P}_{0}\mathbf{U}) = \operatorname{tr}(\mathbf{P}_{0}\mathbf{U}\mathbf{U}^{\dagger}) = \operatorname{tr}(\mathbf{P}_{0}) = n$

• How should we parameterize the unitary transform?

Products of 2-by

- Cayley form?
- Exponential of an antisymmetric matrix?













Matrix polynomial evaluation

- Clearly the rate-determining step in Chebyshev approach – Is it done done with optimal efficiency? NO!
- Consider a polynomial of order p:

$$f(X) = a_0 + a_1 \mathbf{X} + a_2 \mathbf{X}^2 + a_3 \mathbf{X}^3 + \dots + a_p \mathbf{X}^p$$

- Simple evaluation requires *p*-1 matrix multiplies because all powers of **X** are required. Build up powers order by order, and increment the sum.
- Let's focus on a specific example: p=24 (23 multiplies)

 $f(X) = a_0 + a_1 \mathbf{X} + a_2 \mathbf{X}^2 + a_3 \mathbf{X}^3 + \dots + a_{24} \mathbf{X}^{24}$



Subdivision of the polynomial by a number approaching √p...
 Arrange the matrix powers in Homer form for greatest efficiency

$$f(\mathbf{X}) = a_0 + (a_1\mathbf{X} + a_2\mathbf{X}^2 + a_3\mathbf{X}^3 + a_4\mathbf{X}^4 + a_5\mathbf{X}^5)$$

- + $\mathbf{X}^{5}(a_{6}\mathbf{X}+a_{7}\mathbf{X}^{2}+a_{8}\mathbf{X}^{3}+a_{9}\mathbf{X}^{4}+a_{10}\mathbf{X}^{5}$
- +**X**⁵ $(a_{11}$ **X** + a_{12} **X**² + a_{13} **X**³ + a_{14} **X**⁴ + a_{15} **X**⁵
- + $\mathbf{X}^{5}(a_{16}\mathbf{X} + a_{17}\mathbf{X}^{2} + a_{18}\mathbf{X}^{3} + a_{19}\mathbf{X}^{4} + a_{20}\mathbf{X}^{5}$
- + $\mathbf{X}^{5}(a_{21}\mathbf{X} + a_{22}\mathbf{X}^{2} + a_{23}\mathbf{X}^{3} + a_{24}\mathbf{X}^{4}))))$
- 8 multiplies if I count right... still same number of additions
- This algorithm was (first) discovered by Paterson & Stockmeyer

M	atrix multipl	ies and saved	l intermediat	es
- p is $- M is$ $- S is$	the degree of polyno number of matrix n the number of saved	omial. nultiplies. Savings to matrices. The num	ends to $\frac{1}{2}\sqrt{p}$ ber tends to \sqrt{p}	
	р	S	М	
	128	10	21	
	256	15	30	
	336	17	35	
	1024	31	62	
	1600	39	78	

CPU tim for a den	es (in sec.) for e se matrix of side	valuating matrix pol e length 422	ynomials of degree
Γ	p	t(conventional)	t(fast)
	180	23	4
	256	33	5
	480	62	8
	1024	133	13

Chebyshev polynomials

• Order p Chebyshev polynomial is a polynomial of degree p. The polynomials are defined by:

$$T_0(X) = T$$

 $T_1(X) = X$
 $T_2(X) = 2X^2$

 $\mathbf{T}_{n+m}(\mathbf{X}) = 2\mathbf{T}_n(\mathbf{X})\mathbf{T}_m(\mathbf{X}) - \mathbf{T}_{|n-m|}(\mathbf{X})$

- 1

 Our polynomial results generalize to Cheyshev polynomials. The only difference is that the coefficients change in the resummed series, by simple recurrences.
 See J. Chem. Phys. 119, 4117 (2003).





Poly	Polynomial degrees and matrix multiplications				
• Test ex – Erro – <i>p</i> is	ample is C ₆₀ r is the deviation degree of poly	H ₁₂₂ , BLYP/ on in the energy nomial, <i>M</i> is nu	/STO-3G (a.u.) vs diago umber of matrix		
	function	р	М	error	
	CEF	129	21	1.2×10-6	
	FD	269	31	1.1×10-6	
	tanh	269	31	1.1×10-6	
	exp	412	39	2.8×10-6	

Test ex – Erro – p is	cample is C ₆₀ or is the deviation degree of polym	H ₁₂₂ , BLYP, on in the energy nomial, <i>M</i> is no	/STO-3G / (a.u.) vs diag umber of matr	onalization. ix multiplies
	function	р	М	error
	CEF	129	21	1.2×10 ⁻⁶
	FD	269	31	1.1×10 ⁻⁶
	tanh	269	31	1.1×10 ⁻⁶
	exp	412	39	2.8×10-6

Comparison of methods: 2-d water clusters			er clusters
CPU times (sec waters via CEF) for water clu versus diagon	sters made of cells alization.	s each with 8
size	basis	t(CEF)	t(diag)
(H ₂ 0) ₇₂	STO-3G	3.4	1.4
(H ₂ O) ₁₂₈	STO-3G	10.5	11.0
(H ₂ O) ₂₀₀	STO-3G	21.1	40.1
(H ₂ O) ₂₈₈	STO-3G	42.4	91.7
(H ₂ O) ₇₂	6-31G**	211	111
(H ₂ O) ₁₂₈	6-31G**	674	802

Numbers of matrix multiplies: 2-d systems

- Water clusters with STO-3G:
- Water clusters with 6-31G**
- Compare against about 10 or so for dense matrix multiplication.

Outline

- 1. Overview of density functional theory
- 2. Fast methods for Fock matrix formation
- 3. The diagonalization problem
- 4. Excited states and time-dependent DFT
 - A sketch of the theoryCharge-transfer excited states

Wavefunction picture of ground and excited states

• Ground state:

- Zero order picture is mean-field theory... electrons paired up in molecular orbitals
- Brillouin theorem: HF determinant is uncoupled from single excitations. Leading ground state correction is doubles.
- Excited states:
 - Hamiltonian-uncoupled from the ground state
 Mathematically this is single excitation CI (CIS).

Time-Dependent Density Functional Theory for Excited States.

- Hohenberg-Kohn and Kohn-Sham theorems
- First numerical tests on closed shell molecules
 - Jamorski, Casida, Salahub, JCP, 104, 5134 (1997)
- Limitations for Ryberg states, higher excited states
 - Casida, Salahub et al, JCP 108 4439 (1998)



- Begin with the Kohn-Sham equation of motion: $\begin{bmatrix} \hat{F} & \hat{\sigma} \end{bmatrix} = i \frac{\partial P}{\partial r}$
- Apply time dependent perturbation, and look for natural resonances in the linear response. The result is: $\begin{bmatrix} \mathbf{A} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{x} \end{bmatrix} = \begin{bmatrix} \mathbf{B} & \mathbf{x} \end{bmatrix}$

$$\begin{bmatrix} \mathbf{x} \\ \mathbf{x} \end{bmatrix} = \boldsymbol{\omega} \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} \mathbf{z} \\ \mathbf{y} \end{bmatrix}$$

- TDDFT looks like RPA with a dressed response matrix, and the dressing is *formally exact*.
- Tamm-Dancoff approximation (B=0) simplifies TDDFT to have the form of CIS with dressed response:
 So Hirata, MHG, CPL <u>314</u>, 291-299 (1999)



• Numerical tests are essential to assess the validity of TDDFT within these approximations.



• By contrast, the TDDFT in the adiabatic approximation yields a very simple dressing: no energy dependence...

$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	State	CIS	TDDFT	CASPT2	Expt.
$ \begin{array}{cccccccccccccccccccccccccccccccccccc$	${}^{2}B_{1u}$	2.20	1.05	0.99	0.73
$ \begin{array}{c ccccccccccccccccccccccccccccccccccc$	${}^{2}B_{2g}$	2.49	2.16	1.89	1.9
	${}^{2}B_{3g}$	3.53	2.78	2.70	2.7
	$^{2}A_{g}$	5.71	2.92		(2.7)
	${}^{2}\mathbf{B}_{1g}$	5.70	2.97		(2.7)
${}^{2}B_{2g}$ 4.25 3.75 3.98 4.02 ${}^{2}B_{2u}$ 3.87 ${}^{2}B_{1u}$ 4.60 4.19 4.03	${}^{2}B_{3g}$	3.83	3.53	3.24	3.25
${}^{2}B_{2u}$ 3.87 ${}^{2}B_{1u}$ 4.60 4.19 4.03	${}^{2}B_{2g}$	4.25	3.75	3.98	4.02
² B ₁ , 4.60 4.19 4.03	${}^{2}B_{2u}$		3.87		
	${}^{2}B_{1u}$	4.60	4.19	4.03	









Why is the Coulomb attraction missing?

- Present-day TDDFT suffers from "self-interaction" error.
 An electron can artificially see itself
- Exact exchange is essential to get the Coulomb interaction associated with charge transfer...

• CIS $K_{ia,ia} = -(ii)aa$

• TDDFT
$$K_{ia,ia} = \iint d\mathbf{r}_i d\mathbf{r}_j i(1) a(1) \kappa(\mathbf{r}_i, \mathbf{r}_j) i(2) a(2)$$

• The kernel is local in standard density functionals, so this vanishes when *i* and *a* do not overlap...

A.Dreuw, J.Weisman, MHG, JCP 119, 2943 (2003), JACS 126, 4007 (2004)

Why are the charge transfer states much too low?

- In Hartree-Fock theory, $(\varepsilon_a \varepsilon_i)$ is an estimate of the CT excitation energy
- Unlike Hartree-Fock theory, the same potential is used to obtain occupied and virtual eigenvalues in DFT.
 LUMO is hence more strongly bound than by Hartree-Fock
- This is a self-interaction error which is not correctly cancelled by the exchange potential.

A work-around to calculate charge transfer states

- CIS is free of the self-interaction problem... but does not properly describe correlation and relaxation...
- A hybrid CIS/DFT approach for charge-transfer states.
 - Get lowest CT state at large separation by ground state DFT
 - Compute the same state at the same separation (r_0) by CIS.
 - The difference defines a "correlation+relaxation" offset, by which the full CIS potential curve can be shifted.

 $\boldsymbol{\omega}_{CT}(r) = \boldsymbol{\omega}_{CT}^{CIS}(r) + \left[\boldsymbol{\omega}_{CT}^{DFT}(r_0) - \boldsymbol{\omega}_{CT}^{CIS}(r_0) \right]$

Combine with regular TDDFT for intramolecular excitations
 A.Dreuw, MHG, JCP 119, 2943 (2003), JACS 126, 4007 (2004)

Excited states via the hybrid TDDFT/CIS approach 45 3.5 2.5 Q states 1.5 ∆KS–DFT offset 0.5 ground state -0.5 4 5 7 8 9 10 6

Non-photochemical quenching (NPQ)

- Non-photochemical quenching (NPQ) is the partial shutdown of light-harvesting under high-light conditions in green plants.
- Correlates with production of the carotenoid zeaxanthin (Zea) from violaxanthin via the xanthophyll cycle.
- What is the mechanism of energy relaxation from an excited chlorophyll (Chl) molecule?
- Could a dimer between Zea and Chl be involved?
- · Calculations in collaboration with Graham Fleming's group.

Carotenoid-chlorophyll interactions • Zeaxanthin 11 conjugated bonds • Violaxanthin 9 conjugated bonds • Chlorophyll a (simple model) H = H = H = H



TDDFT again fails for charge transfer (CT) states

- TDDFT/TDA BLYP/3-21G calculations
- Lowest CT state is unphysically low (error \longrightarrow 3 eV) (We calculate IP(zea) = 4.96 eV, EA(chl) = -1.25 eV)
- CT states do not show Coulomb attraction at long distances !!











Conclusions

- Continued testing and development of TDDFT is needed
 Charge-transfer states fail with standard functionals
- Charge-transfer states may play a significant role in NPQ
 An experiment looking for transient absorption from the Zea radical cation is a signature of the dimer...
 - This appears to require complex formation- relatively close approach
 of zea and chl in the favored orientation.
- Resonant energy transfer between Chl and Zea is also possible. Differences between Zea and Vio account for NPQ.

Q-Chem 2.0: A review of features and capabilities

J.Kong, C.A.White, A.I.Krylov, C.D.Sherrill, R.D.Adamson, T.R.Furlani, M.S.Lee, A.M.Lee, S.R.Gwaltney, T.R.Adams, C.Ochsenfeld, A.T.B.Gilbert, G.S.Kedziora, V.A.Rassolov, D.R.Maurice, N.Nair, Y.Shao, N.A.Besley, P.E.Maslen, J.P.Dombroski, H.Daschel, W.Zhang, P.P.Korambath, J.Baker, E.F.C.Byrd, T.Van Voorhis, M.Oumi, S.Hirata, C.-P.Heu, N.Ishikawa, L.Elozion, A.Warshel, B.G. Johnson, P.M.W.Gill, C. P. Heu, N.Ishikawa, L.Elozion, A.Warshel, B.G. Johnson, P.M.W.Gill, C. P. Heu, N.Ishikawa, L.Elozion, A.Warshel, B.G. Johnson, P.M.W.Gill,

C.-P.Hsu, N.Ishikawa, J.Florian, A.Warshel, B.G.Johnson, P.M.W.Gill, M.Head-Gordon, J.A.Pople.

Journal of Computational Chemistry (2000) **21**, 1532 (Special issue on ab initio methods for large molecules)