

THE UNIVERSITY of TEXAS

HEALTH SCIENCE CENTER AT HOUSTON School of Health Information Sciences

Quantum Chemistry II

For students of HI 6327 "Biomolecular Modeling"

Jochen Heyd, Ph.D., and Willy Wriggers, Ph.D. School of Health Information Sciences

http://biomachina.org/courses/modeling/12.html

Overview

- Car-Parrinello (brief)
- QM/MM embedding
- The HSE DFT functional
- Accuracy of various QC methods
- Calculating molecular properties
- Force field parameterization
- A case study

Car-Parrinello MD

- CPMD is a DFT method based on the following idea:
- Small perturbations in the positions of the nuclei only produce a small change in the solution of the SCF procedure.
- → Evolve atom positions and the wave function simultaneously



© Axel Kohlmeyer

Quantum Mechanics/Molecular Mechanics (QM/MM)

Goal: Investigate chemical reactions in condensed phase environments.

Exploit explicit treatment of electrons with QM approach in combination with computationally inexpensive MM approach to treat condensed phase environment.



Embedding Approaches



- Steric embedding
- Charge embedding
- Link atom
- Covalent embedding

Link Atom Approach

- 1) Include hydrogen to fill valence (or fluorine or methyl).
- 2) Maintain QM to MM covalent terms (bonds, angles, and dihedrals)
- 3) QM sees all MM atoms (as charges) and vice versa. Omission of link atom to MM atom interactions leads to poor results due to polarization of link atom - carbon bond. This is caused by the carbon "feeling" the MM field while the link atom does not.



Apply constraints to keep link atom along the C-CH bond.

© Alex MacKerell

Link-Atom Embedding: Problems

and the second second

F	Potential Problems with Link Atom Idea
	Extra degrees of freedom which somehow need to
	be removed; i.e. the link atom somehow needs to be connected to the MM part of the simulation
	Electronic structure at boundary will be very
	different if H and the atom it replaces do not have
ad dal sait dal dal dal dal dal dan sa	similar electronegativities

Selection of QM Region

• In solution: Reactants being treated by QM are surrounded by MM waters: no covalent bonds between QM atoms and MM atoms

• Enzyme active site: Necessity of including amino acid or coenzyme functional groups in QM region requires treatment of covalent link between QM and MM regions (frontier region).

- Valences of the frontier atoms of QM region must be filled: Link Atoms Frontier Orbitals
- Other guidelines:

1) Perform test calculations using different QM/MM partitions to see if the charge distribution reproduces full QM values in the gas phase.

2) ALWAYS place QM/MM frontier (link atom or orbital) at an aliphatic carbon as it is desirable to have the smallest possible charge distribution at the frontier.

Questions: QM/MM?

Next: Density Functional Theory (cont.)

Density Functional Theory

• Hohenberg and Kohn (1964)

$$E = E[\rho]$$

• The energy is composed of several parts

$$E_{tot} = T_e + V_{ne} + J_{ee} + E_{XC}$$

Exchange and Correlation Functionals

$$E_{XC} = E_{XC}[\rho]$$

$$E_{XC} = E_{XC}[\rho, \nabla \rho]$$

Local spin density approximation (LSDA)

Generalized gradient approximation (GGA)

BLYP, PBE, ...

$$E_{XC} = E_{XC}[\rho, \nabla \rho, \tau]$$

Meta-generalized gradient approximation (mGGA)

VSXC, TPSS, ...

Exchange Functionals and Hartree-Fock Exchange

- DFT exchange functionals perform well for systems with slowly varying densities
- HF exchange is exact but difficult to pair with a DFT correlation functional
- Hybrid exchange functionals mix DFT and HF
- Hybrids can describe systems with rapid variations in electron density

Why Do We need Exact (HF) Exchange?



Why not always use hybrids?

- Pure DFT scales linearly with system size
- HF exchange is computationally expensive
- Linear scaling methods for evaluating exact exchange depend strongly on the HOMO-LUMO (or band) gap
- Truncation schemes often lead to errors of unknown magnitude and can complicate the selfconsistent field convergence

Infinite Carbon Atom Chain



The Central Idea

Let the exchange potential delocalize, but only among near neighbors!

Partitioning the Coulomb Operator

The Coulomb operator can be separated as follows:



The parameter $\boldsymbol{\omega}$ controls the splitting into short and long range.

Partitioning the Coulomb Operator



Distance

Infinite Carbon Atom Chain



(6,6) Carbon Nanotube: Decay of the Exchange Interaction



Partitioning the Coulomb Operator



Distance

Heyd-Scuseria-Ernzerhof Hybrid Functional

Starting from the standard PBE0 hybrid

$$E_{XC}^{PBE0} = a \cdot E_X^{HF} + (1-a) \cdot E_X^{PBE} + E_C^{PBE} \qquad a = \frac{1}{4}$$

$$E_X = a \cdot E_X^{HF} + (1-a) \cdot E_X^{DFT}$$

$$E_X = a \cdot E_X^{HF,SR}(\omega) + a \cdot E_X^{HF,LR}(\omega) + (1-a) \cdot E_X^{DFT,SR}(\omega)$$

$$+ E_X^{DFT,LR}(\omega) - a \cdot E_X^{DFT,LR}(\omega) \qquad \text{assumed to cancel}$$

$$E_{XC}^{HSE} = a \cdot E_X^{HF,SR}(\omega) + (1-a) \cdot E_X^{\omega PBE,SR}(\omega)$$

$$+ E_X^{\omega PBE,LR}(\omega) + E_C^{PBE}$$

$$a = \frac{1}{4}$$

J. Chem. Phys. 118, 8207 (2003)

Questions: HSE?

Next: DFT performance on various systems

Atomization Energies: Basis set dependence for a set of 55 molecules

Basis Set	PBE	PBE0	HSE
3-21G	12.94	22.56	17.91
6-31G*	7.23	6.94	6.40
6-311G**	6.14	6.47	4.88
6-311++G(3df,3pd)	5.92	3.49	3.38

Atomization Energy MAE in kcal/mol

Enthalpies of Formation: G3 Test Set (223 Molecules)



Optimized Bond Lengths: T-96R Test Set (96 Molecules)



Lattice Constants



Band Gaps (Semi-Conductors)



Time Per Complete SCF



Questions: DFT?

Next: Calculating Molecular Properties

Calculating Molecular Properties

- All molecular properties are derivatives of either the energy or the electronic density / wavefunction.
- Molecular geometry
- Vibrational frequencies
- Electron density and HOMO/LUMO
- Partial charges
- Electrostatic potential

Potential Energy Surfaces

- Equilibrium molecular structures correspond to the positions of the minima in the valleys on a PES
- Energetics of reactions can be calculated from the energies of the minima for reactants and products
- A reaction path connects reactants and products through a mountain pass
- A transition structure is the highest point on the lowest energy path
- Reaction rates can be obtained from the height and profile of the potential energy surface around the transition structure

Important Points

© Todd J. Martinez

Calculation of Harmonic Vibrational Frequencies

Harmonic vibrational frequencies can be calculated via the second derivative of the energy with respect to the atomic coordinates.

Important uses:

- Characterization of transition states
- Calculation of IR and Raman spectra

Intrinsic Reaction Coordinate (IRC)

Electron Density, HOMO and LUMO

- The electron density is the probability distribution of finding an electron at a specific point in space.
- The highest occupied and lowest unoccupied molecular orbitals (HOMO and LUMO) are important tools to understand molecular reactivity.

Population Analysis

• Population analysis methods partition the electron density between the nuclei so that each nucleus has a number of electrons associated with it.

• There is no quantum mechanical operator for the atomic charge and so any partitioning scheme must be arbitrary.

 Many method have been devised, such as: Mulliken and Löwdin analysis, Bader's theory of atoms in molecules, natural population analysis

• Mulliken method: All the electron density ($P_{\mu\mu}$) in an orbital is allocated to the atom on which ϕ_{μ} is located.

• For each element $\phi_{\mu}\phi_{\nu}$ of the density matrix, half of the density is assigned to the atom on which ϕ_{μ} is located and half to the atom on which is located

• The net charge on atom A is calculated by subtracting the number of electrons from the nuclear charge, Z_A

$$q_{A} = Z_{A} - \sum_{\mu=1;\mu \text{ on } A}^{K} P_{\mu\mu} + \sum_{\mu=1;\mu \text{ on } A}^{K} \sum_{\nu=1;\nu\neq\mu}^{K} P_{\mu\nu} S_{\mu\nu}$$

Partial Atomic Charges

* "Atoms" are not well-defined in molecules – there is no quantum mechanical operator corresponding to an atom. This leads to ambiguity in the definition of an atomic charge Population Analysis Schemes Basically, sum over all electrons using the basis functions of a given atom Depends on the atom-centered nature of the basis set Breaks down as the basis functions become more delocalized - results do not usually converge with increasing basis set!

Atoms in Molecules

Dipole Moment

- The electric multipoles reflect the distribution of charge in a molecules, i.e. it is a measure of the asymmetry in the molecular charge distribution
- The simplest electric moment is the dipole.
- The dipole moment of a distribution of charges q_i located at positions l_i is given by $\sum_i q_i l_i$
- The dipole moment of a complex molecule can be represented approximately in the form of the vectorial sum of the moments belonging to the individual bonds.

$$p = ql$$

• The quadrupole is the next electric moment. A molecule has a non-zero electric quadrupole moment when there is a non-spherically symmetrical distribution of charge.

Electrostatic Potential

- The electrostatic potential at point $r, \phi(r)$, is defined as the work done to bring unit positive charge from infinity to the point.
- the electrostatic interaction energy between a point charge q located at r and the molecule equals $q\phi(r)$
- The electrostatic potential *V*(**r**) that is produced at any point **r** by the electrons and nuclei (A) of the system is given by

$$V(r) = \sum_{A} \frac{Z_A}{\left|r - R_A\right|} - \int \frac{dr'\rho(r)}{\left|r' - r\right|}$$

- The electrostatic potential is particularly useful for rationalising the interactions between molecules and molecular recognition processes. The electrostatic force is primarily responsible for long-range interactions between molecules.
- Electrostatic potential contours can be used to propose where electrophilic attack might occur; electrophiles are often attracted to regions where the electrostatic potential is most negative.

Electrostatic Potential

© A.R. Leach

© A.R. Leach

Electrostatic potential mapped onto the electron density surface of formamide. Red indicates negative electrostatic potential and blue is positive potential.

Naphthalene MEP © Gaussian Inc.

Electrostatic potential contours around cytosine. Negative contours are dashed, the zero contour is bold. The minima near N3 and O are marked.

© A.R. Leach

http://educ.gaussian.com/visual/ESP/html/ESPGaussView.htm

Questions: Calculating Molecular Properties?

Next: Parametrizing Potentials

Force-Field Parameterization

• These parameters have to be chosen against some model of *reality*. Traditionally this is taken from experimental data. For other classes of molecules data is more scarce. Even when experimental data is available, some force field parameters can be hard to determine from it.

• Bond lengths and angles can be determined from crystallography. Van der Waals can be determined from crystal structures. However, torsional potentials can be hard to determine, as are atomic charges.

• Recently some force fields have been parameterized using data from electronic structure calculations. This allows the determination of force field parameters for molecules where little or no data exists.

• It also allows the determination of parameters in a consistent fashion. Whereas experimental data relevant to one set of parameters may be available in the gas phase, while that for another may be available for a solid, parameters from electronic structure calculations can all be found from the same phase (usually the gas phase). This can be important as there can be large differences in molecular structure in different phases.

• Full torsional potentials can be determined from electronic structure calculations. Partial atomic charges can also be found from calculations. However, there are several different methods to find charges, with little agreement between them.

Torsional Potential Terms

 $\mathsf{E}_{\mathsf{tors}} = \Sigma \left\{ \mathsf{V}_{\mathsf{i}} \cos[(\theta - \theta_0)] + \mathsf{V'}_{\mathsf{i}} \cos[2(\theta - \theta_0)] + \mathsf{V''}_{\mathsf{i}} \cos[3(\theta - \theta_0)] \right\}$

©H. B. Schlegel

Di-Methyl-Sulfonamide (DMSA): A Case Study

- Sulfonamides are important in medicine for treating infections
- A medical company needed force field parameters
- How do we proceed?

DMSA: Computational Approach

- 1. Optimize the geometry
- 2. Scan the potential energy surface
- 3. Optimize transition states
- 4. Characterize transition states
- 5. Perform IRC run on transition states
- 6. Fit a torsional potential for use with the force field

DMSA: The Full Story

Torsion angle between methyl-groups

DMSA: Comparison of Methods

	GS energy in HT's	ΔE of TS in $\frac{k c a l}{m o l}$	$\mathcal{L}_{CNSC,GS}$	$\mathcal{L}_{CNSC,TS}$	$\angle_{HNSC,TS}$
HF/CEP -31G	not comp.	6.58	-104.62°	179.39°	-0.61°
$\mathrm{HF}/3-21\mathrm{G}$	-677.74974	15.94	-102.56°	170.40°	$+1.22^{\circ}$
$HF/6-31G^{*}$	-681.41657	7.39	-98.93°	135.63°	-4.69°
HF/6-311G**	-681.51363	7.41	-95.89°	136.68°	-4.06°
HF/6-311+G(2d,p)	-681.56260	7.41	-97.50°	133.86°	-2.66°
HF/6-311+G(2df,p)	-681.59221	7.16	-96.50°	139.45°	-3.70°
MP2/6-31G*	-682.34172	7.10	-98.23°	127.45°	-4.25°

J. Heyd, W. Thiel, and W. Weber, "Rotation and inversion barriers in N- methylmethanesulfonamide from ab initio calculations", Theochem - J. Mol. Struc. 391, 125 (1997)

DMSA: The Full Story

Resources and Further Reading

WWW:

-Computational Chemistry List: http://www.ccl.net

-http://educ.gaussian.com/visual/index.htm

Books:

• "Molecular Modeling-Principle and Application", Andrew R. Leach, Chapter 2, 3

• "Exploring Chemistry with Electronic Structure Methods", James B. Foresman, Eleen Frisch, Gaussian Inc.

• "Ab initio Molecular Orbital Theory", W.J. Hehre, L. Radom, J.A. Pople, P.V.R. Schleyer

• "Combined Quantum Mechanical and Molecular Mechanical Methods", J. Gao, M. A. Thompson (ed.), ACS Symposium Series no. 712, American Chemical Society, Washington 1998

• "Computer Modeling of Chemical Reactions in Enzymes and Solutions", A. Warshel, John Wiley & Sons, New York 1991.

• "A Chemist's Guide to Density Functional Theory", W. Koch & M. C. Holthausen, WILEY-VCH, 2000.

Papers:

Quantum Chemical Methods for Large Molecules, Journal of Computational Chemistry, vol. 21, John Wiley & Sons, New York 2000

Acknowledgements:

- Matt Jacobson, University of San Francisco
- Alexander D. MacKerell Jr., University of Maryland at Baltimore
- Todd J. Martinez, University of Illinois at Urbana-Champaign
- G. Matthias Ullmann, Uni Heidelberg
- H. B. Schlegel, Wayne State University

Additional Slides

Calculation of Harmonic Vibrational Frequencies

© Gaussian Inc.

The total energy of a molecule comprising N atoms near its equilibrium structure may be written as

$$E = T + V = \frac{1}{2} \sum_{i=1}^{3N} \dot{q}_i^2 + V_{eq} + \frac{1}{2} \sum_{i=1}^{3N} \sum_{j=1}^{3N} \left(\frac{\partial^2 V}{\partial q_i \partial q_j} \right)_{eq} q_i q_j.$$
(1)

Here, the mass-weighted cartesian displacements, q_i , are defined in terms of the locations x_i of the nuclei relative to their equilibrium postiions $x_{i,eq}$ and their masses M_i ,

$$q_i = M_i^{1/2}(x_i - x_{i,eq})$$

 V_{eq} is the potential energy at the equilibrium nuclear configuration, and the expansion (1) of the vibrational energy in terms of a power series is truncated at second order [16]. For such a system, the classical-mechanical equations of motion take the form

$$\ddot{q}_{j} = -\sum_{i=1}^{3N} f_{ij}q_{i}$$
 $j = 1, 2, ..., 3N.$ (2)

© J.A. Pople

Calculation of Harmonic Vibrational Frequencies

The f_{ij} , termed quadratic force constants, are the second derivatives of the potential energy with respect to mass-weighted cartesian displacements, evaluated at the equilibrium nuclear configuration, that is,

$$f_{ij} = \left(\frac{\partial^2 V}{\partial q_i \partial q_j}\right)_{eq}$$

The f_{ij} may be evaluated by numerical second differentiation,

$$\frac{\partial^2 V}{\partial q_i \partial q_j} \simeq \frac{\Delta(\Delta V)}{\Delta q_i \Delta q_j}$$

by numerical first differentiation of analytical first derivatives,

$$\frac{\partial^2 V}{\partial q_i \partial q_j} \simeq \frac{\Delta(\partial V/\partial q_j)}{\Delta q_i}$$

Equation (2) may be solved by standard methods [16] to yield a set of 3N normal-mode vibrational frequencies. Six of these (five for linear molecules) will be zero as they correspond to translational and rotational (rather than vibrational) degrees of freedom.[‡]

[16] E.B. Wilson, J.C.Decius, and P.C. Cross, Molecular Vibrations, McGraw-Hill, New York, 1955.

© J.A. Pople

Total Electron Density Distribution

The integral of $\rho(r)$ over all space equals the number of electrons in the system:

Isometric projection, height above the plane: magnitude of the electron density

formamide

Surface, contour 0.0001 au

HOMO highest occupied molecular orbital

LUMO lowest unoccupied molecular orbital © A.R. Leach

Molecular Orbital & Electron Density

Cutplane at XY-plane of aspirine

Cutplane at XZ-plane of aspirine © CSC

Calculating Partial Atomic Charges

$$\phi_{qm}(\mathbf{r}) = \sum_{j=1}^{Atoms} \frac{Z_j}{|\mathbf{r} - \mathbf{R}_j|} - \int \frac{\rho(\mathbf{r'})}{|\mathbf{r} - \mathbf{r'}|} d\mathbf{r'}$$

NuclearElectronChargesDensity

$$\phi_{\text{classical}}(\mathbf{r}) = \sum_{j=1}^{Atoms} \frac{q_j}{|\mathbf{r} - \mathbf{R}_j|}$$

© G. Matthias Ullmann

Gaussian03(G03) - an *ab initio* Package

- Single Point Energy Calculation
- Geometry Optimization
- Frequency Calculation
- Chemical Reactions and Reactivity (TS, IRC, PES scan)
- Modeling Excited States
- Modeling Systems in Solution

What Properties Can G03 Provide?

- Molecular orbital information
- Population analysis & atomic charges
- IR and Raman spectra
- NMR shielding tensors
- Multipole moments
- Electron affinities and ionization potentials
- Polarizabilities and hyperpolarizabilities

G03 Basis Functions / Methods

SCF methods	HF		Hartree-Fock	
	CASSCF		Complete active space SCF	
Post-SCF	MP2,MP3,N	/IP4	Møller-Plesset Pert. Theo.	
methods:	CI, QCISD		Configuration Interaction	
	CCSD, QCI	SD	Couple cluster theory	
DFT	BLYP, PW91	B3LYP,	Backe-style 3-Parameter DFT (Lee-Yang-Parr correlation functional)	
High Accuracy	G1, G2,		Gaussian1, Gaussian2	
Energy	CBS-		Complete Basis Set Methods	
Excited State	CIS			
Semi-empirical CNDO, MINDO/3,				
	AM1, PM3	-		

Input File Example

Spin Multiplicity

The spin multiplicity M controls the spin state of your molecule.

M	spin state	
1	singlet	\longrightarrow No unpaired electrons (e.g. closed shell)
2	doublet	—> One unpaired electron
3	triplet	—> Two unpaired electrons of like spin
4	quartet	Three unpaired electrons of like spin

The spin multiplicity can be calculated from the quantum number S of the total spin or from the number n_u of unpaired electrons:

 $M = 2 S + 1 = n_u + 1$

Internal Coordinates (Z-matrix)

	С
1	C 1 r1
2	H 1 r2 2 a
3	H 1 r2 2 a 3 180.
4	H 2 r2 1 a 3 0.
5	H 2 r2 1 a 4 0.
6	
	r1 1.31
	r2 1.07
	a 121.5

Å, degree

Population Analysis & Atomic Charge

Condensed to atoms (all electrons):

		1	2	3	4	5	6
1	С	4.780096	0.612484	0.393153	0.393153	-0.025895	-0.025895
2	С	0.612484	4.780096	-0.025895	-0.025895	0.393153	0.393153
3	Η	0.393153	-0.025895	0.596856	-0.024702	-0.005077	0.002118
4	Η	0.393153	-0.025895	-0.024702	0.596856	0.002118	-0.005077
5	Η	-0.025895	0.393153	-0.005077	0.002118	0.596856	-0.024702
6	Η	-0.025895	0.393153	0.002118	-0.005077	-0.024702	0.596856
Total atomic charges:							

- 1 C -0.127095
- 2 C -0.127095
- 3 H 0.063548
- 4 H 0.063548
- 5 H 0.063548
- 6 H 0.063548

Sum of Mulliken charges= 0.00000

Potential Energy Surface Scan

#hf/6-31+g* scan

--PES scan--

0 1 o h 1 0.947 h 1 r 2 106.5

r 1.0 10 -0.01

Potential Energy Surface Scan

N r SCF

1	1.0000	-76.01507
2	0.9900	-76.01595
3	0.9800	-76.01667
4	0.9700	-76.01722
5	0.9600	-76.01758
6	0.9500	-76.01774
7	0.9400	-76.01768
8	0.9300	-76.01739
9	0.9200	-76.01685
10	0.9100	-76.01605
11	0.9000	-76.01497

