

# Výpočetní chemie a počítačové modelování

Jan Řezáč

UOCHB AV ČR

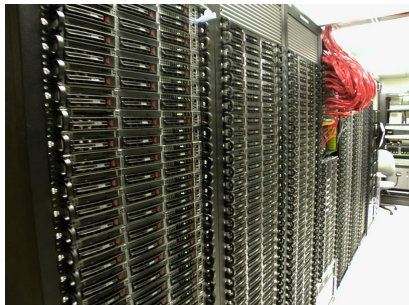
7. října 2021

- Samotná teorie umožňuje odvodit chemické vlastnosti
- Molekuly a jejich systémy plně popsány kvantovou mechanikou

# Proč počítat, když máme laboratoře?

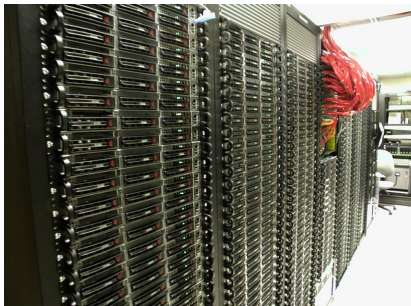
- Studium experimentálně nepřístupných vlastností
- Odhalování strukturní podstaty jevů
- Často levnější než experiment

# Proč experimentovat, když máme počítače?



- Nemáme jich dost.

# Proč experimentovat, když máme počítače?



- Nemáme jich dost.
- Výpočetně náročné
  - Přesné kvantové výpočty  $\sim 100$  atomů
  - Molekulová mechanika  $\sim 100\,000$  atomů, 1 ms

- 1 Kvantová chemie
- 2 Chemické vlastnosti
- 3 Výpočty - praxe
- 4 Molekulová mechanika
- 5 Molekulová dynamika

# Důležité aproximace

Kvantově-mechanické rovnice nelze vyřešit bez aproximací

# Důležité aproximace

Kvantově-mechanické rovnice nelze vyřešit bez aproximací

## 3 kroky k praktickým výpočtům

- Born-Oppenheimerova aproximace
- MO jako lineární kombinace AO
- Jednoelektronová aproximace



# Born-Oppenheimerova aproximace

- Separace vlnové funkce jader a elektronů
- $m_N \gg m_e \rightarrow v_n \ll v_e$
- Řešíme schrödingerovu rovnici pro elektrony v elektrostatickém poli jader

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{ee} + \hat{V}_{Ne}$$

- Dostatečné pro statické vlastnosti molekuly
- Pohyb jader lze často popsat jen klasicky

# Born-Oppenheimerova aproximace

- Separace vlnové funkce jader a elektronů
- $m_N \gg m_e \rightarrow v_n \ll v_e$
- Řešíme schrödingerovu rovnici pro elektrony v elektrostatickém poli jader

$$\hat{H} = \hat{T}_N + \hat{T}_e + \hat{V}_{NN} + \hat{V}_{ee} + \hat{V}_{Ne}$$

- Dostatečné pro statické vlastnosti molekuly
- Pohyb jader lze často popsat jen klasicky

# Reprezentace vlnové funkce

- N elektronů
- Praktické rozepsat jako produkt jednoelektronových funkcí
- Prostý produkt

$$\Psi(x_1, x_2) = \phi_1(x_1)\phi_2(x_2)$$

- Není antisymetrická, porušuje Pauliho princip

$$\Psi(x_1, x_2) = -\Psi(x_2, x_1)$$

- Požadavek splňuje lineární kombinace

$$\Psi(x_1, x_2) = \frac{1}{\sqrt{2}} \{ \phi_1(x_1)\phi_2(x_2) - \phi_1(x_2)\phi_2(x_1) \}$$

# Reprezentace vlnové funkce

- Více elektronů - Slaterův determinant

$$\Psi(x_1, x_2, \dots, x_N) = \frac{1}{\sqrt{N!}} \begin{vmatrix} \phi_1(x_1) & \phi_2(x_1) & \dots \\ \phi_1(x_2) & \phi_2(x_2) & \dots \\ \dots & \dots & \dots \end{vmatrix}$$

# Jednoelektronové vlnové funkce $\phi$

- = molekulové orbitály (MO)

# Jednoelektronové vlnové funkce $\phi$

- = molekulové orbitaly (MO)
- **Lineární kombinace atomových orbitalů (MO-LCAO)**

$$\phi = c_1\chi_1 + c_2\chi_2 + c_3\chi_3 + c_4\chi_4 + \dots$$

- Tvar AO známe z řešení Schr. rovnice pro atom
- Použití více AO než je nutné zpřesní výpočet - báze AO

# Formulace problému

- Řešíme bezčasovou Schrödingerovu rovnici

$$\hat{H}\Psi(r) = E\Psi(r)$$

- Vlnová funkce je určena rozvojovými koeficienty  $c_i(x_j)$

# Hartree-Fockova metoda

- Řešení uvažující skutečnou interakci elektronů mezi sebou je příliš složité

$$\hat{H} = \hat{T}_e + \hat{V}_{ee} + \hat{V}_{Ne}$$

- Rozdělíme problém na jednoelektronové části - **jednoelektronová aproximace**
- Na každý elektron působí průměrné pole ostatních elektronů, (potřebujeme elektronovou hustotu v prostoru, tu získáme z vlnové funkce)
- Veličiny na sobě vzájemně závisí - iterativní řešení (SCF)
- Optimalizujeme koeficienty  $c_i(x_j)$



# Variační princip

## Variační princip

Vlnové funkci základního stavu odpovídá nejnižší dosažitelná energie

- Každý jiný stav (jiná vlnová fce.) má energii vyšší

# Variační princip

## Variační princip

Vlnové funkci základního stavu odpovídá nejnižší dosažitelná energie

- Každý jiný stav (jiná vlnová fce.) má energii vyšší
- Určuje směr kterým vlnovou fci. optimalizujeme
- Optimalizace je ukončena když dosáhneme minima

# Hartree-Fockova metoda

- Získáme vlnovou funkci i energii
- Problém: jednoelektronová aproximace nepopisuje korelaci mezi pohybem elektronů - korelační energie
- několik set atomů

# Post-HF metody

- Dodatečné korekce přidávající korelační E
  - delší vzdálenosti – Londonova disperze
  - krátké vzdálenosti – repulze ("correlation cusp / exchange hole")
- Fluktuační okolo základního stavu jsou popsitelné excitovanou vln. funkcí
- Korelované metody přimíchávají excitace do vlnové funkce
- Je třeba použít velkou bázi AO

# Post-HF metody

- Møller-Plesetova poruchová teorie (MP2, MP3, ...)
- Konfigurační interakce (CI) - rozvoj více Slaterových determinantů, excitace
- Coupled Clusters (CC) - vln. fce rozvoj s excitačními operátory
- Náročnější - desítky atomů

# Semiempirické metody

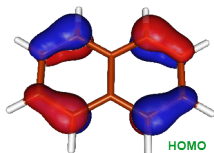
- Zjednodušení HF metody
  - jen valenční elektrony
  - zanedbává víceelektronové integrály
  - přidání mezijádrový potenciál
- Parametrizované na experiment / lepší výpočty
- tisíce atomů

# Teorie funkcionálu hustoty

- Hohenberg-Kohnův teorém: Vlastnosti systému jsou jednoznačně určeny distribucí elektronové hustoty
- Výpočty jednodušší než v případě vlnové funkce
- "Správný" funkcionál ale neexistuje
- Rychlejší a lepší než HF

# Vlastnosti odvoditelné z vlnové funkce

- informace o orbitalech
- dipólový moment a další
- parciální náboje (populační analýza nebo fit)





# Srovnávání energií

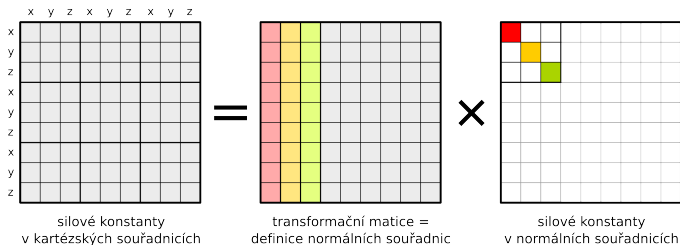
- Výpočet energie umožňuje srovnávat stavy, např. konformace, reakce...

# První derivace E

- = síla působící na jádra
- optimalizace geometrie (v rovnovážné geometrii 1. derivace nulová)
- dynamika - vývoj v čase

# Druhá derivace E

- = silové konstanty
- Převedením do normálních souřadnic získáme vibrační frekvence
- Vibrační partiční funkce  $\rightarrow$  termodynamické vlastnosti



# Excitované stavy

- Lze počítat i excitované stavy (ale není to úplně jednoduché)
- Fotochemie

# Reprezentace molekul v počítači

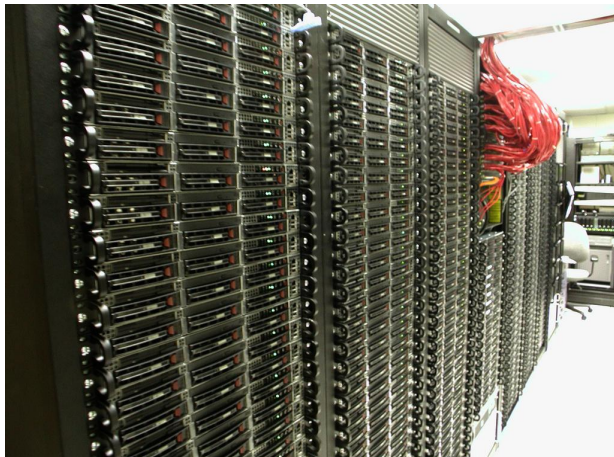
- Vstup: geometrie molekuly
- Lze postavit
- Databáze krystalových struktur (CCDB, PDB/NDB)

```
6
O  -1.551878032  -0.035046733  0.000000434
H  -1.872753946   0.923837822  -0.000000491
H  -0.534611847  -0.101434724   0.000000275
O   1.366706102  -0.089140974  -0.000000160
H   1.888268670  -0.298608582  -0.843170021
H   1.888269052  -0.298606808   0.843169963
~
~
~
8,1      All
```

# Superpočítače a klastry

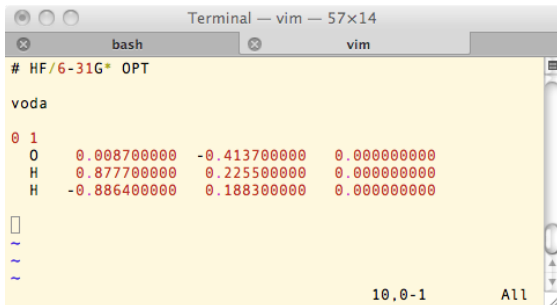
- Výpočty náročné, je jich potřeba hodně
- Paralelizace
- Často se používají klastry běžných PC - levnější než specializované superpočítače
- Užitečné výpočty lze ale provádět i na PC

# Superpočítače a klastry



# Software

- Nejrozšířenější Gaussian (komerční)
- existují další



The screenshot shows a terminal window titled "Terminal — vim — 57x14" with two tabs: "bash" and "vim". The vim editor displays the following text:

```
# HF/6-31G* OPT  
  
voda  
  
0 1  
O 0.008700000 -0.413700000 0.000000000  
H 0.877700000 0.225500000 0.000000000  
H -0.886400000 0.188300000 0.000000000  
  
~  
~  
~
```

At the bottom right of the editor, the coordinates "10,0-1" and "All" are visible.



## Výstup

```

Terminal — bash — 99x38
vim bash
Item      Value      Threshold  Converged?
Maximum Force 0.000077  0.000450  YES
RMS Force     0.000063  0.000300  YES
Maximum Displacement 0.000179  0.001800  YES
RMS Displacement 0.000145  0.001200  YES
Predicted change in Energy=-1.144798D-08
Optimization completed.
-- Stationary point found.

-----
! Optimized Parameters !
! (Angstroms and Degrees) !
-----
! Name Definition      Value      Derivative Info.      !
-----
! R1  R(1,2)              0.9473    -DE/DX = 0.0001      !
! R2  R(1,3)              0.9473    -DE/DX = 0.0001      !
! A1  A(2,1,3)            105.4928  -DE/DX = 0.0         !
-----
GradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGradGrad
-----
Input orientation:
-----
Center   Atomic   Atomic   Coordinates (Angstroms)
Number   Number   Type     X           Y           Z
-----
1         8         0         0.008054   -0.382159   0.000000
2         1         0         0.749788   0.207019   0.000000
3         1         0        -0.757842   0.175241   0.000000
-----
Distance matrix (angstroms):
      1         2         3
1  0  0.000000
2  H  0.947259  0.000000
3  H  0.947255  1.507965  0.000000
Stoichiometry H2O
Framework group CS[5G(H2O)]
Deg. of freedom 3
1003,2 73%

```

## Výstup

```

Terminal — bash — 99x38
vim bash
XXXZ= 0.0000 YYYY= 0.0000 YYYZ= 0.0000 ZZZX= 0
.0000
ZZZY= 0.0000 XXYY= -1.5810 XXZZ= -2.0167 YYZZ= -1
.9085
XXYZ= 0.0000 YYXZ= 0.0000 ZZXY= 0.0000
N-N= 9.289187493137D+00 E-N=-1.990850536058D+02 KE= 7.585992283909D+01
Symmetry A' KE= 7.130832861150D+01
Symmetry A" KE= 4.551594227589D+00
Orbital energies and kinetic energies (alpha):
      1 2
1      0      -20.557867      29.144416
2      0      -1.346171      2.565877
3      0      -0.714291      1.773889
4      0      -0.570823      2.169982
5      0      -0.498217      2.275797
6      V      0.213032      0.965351
7      V      0.306883      0.985211
8      V      1.031695      2.010393
9      V      1.133412      2.477164
10     V      1.168026      3.575933
11     V      1.178554      3.450129
12     V      1.385131      4.133524
13     V      1.431226      2.567415
14     V      2.020530      2.792540
15     V      2.030566      2.800000
16     V      2.067215      2.797459
17     V      2.635703      3.506510
18     V      2.965897      3.937762
19     V      3.977730      9.938350
Total kinetic energy from orbitals= 7.585992283909D+01
Merz-Kollman atomic radii used.

.....

Electrostatic Properties Using The SCF Density

.....
1309,0-1 97%

```

## Výstup

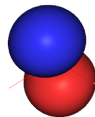
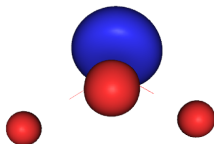
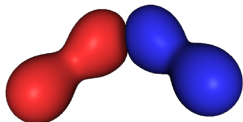
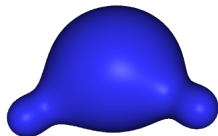
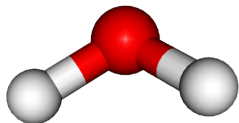
```

Terminal — bash — 99x38
vim bash
Mulliken atomic charges:
  1
  1  O  -0.868746
  2  H   0.434373
  3  H   0.434373
Sum of Mulliken atomic charges =  0.00000
Mulliken charges with hydrogens summed into heavy atoms:
  1
  1  O  0.000000
Sum of Mulliken charges with hydrogens summed into heavy atoms =  0.00000
Electronic spatial extent (au): <R**2>= 18.8102
Charge= 0.0000 electrons
Dipole moment (field-independent basis, Debye):
  X= 0.0000  Y= -2.1990  Z= 0.0000  Tot= 2
.1990
Quadrupole moment (field-independent basis, Debye-Ang):
  XX= -4.1060  YY= -6.0008  ZZ= -7.2056
  XY= 0.0000  XZ= 0.0000  YZ= 0.0000
Traceless Quadrupole moment (field-independent basis, Debye-Ang):
  XX= 1.6648  YY= -0.2300  ZZ= -1.4348
  XY= 0.0000  XZ= 0.0000  YZ= 0.0000
Octapole moment (field-independent basis, Debye-Ang**2):
  XXX= 0.0000  YYY= -1.4305  ZZZ= 0.0000  XYY= 0
.0000
  XXY= -1.3546  XXZ= 0.0000  XZZ= 0.0000  YZZ= -0
.3873
  YYZ= 0.0000  XYZ= 0.0000
Hexadecapole moment (field-independent basis, Debye-Ang**3):
  XXXX= -5.3633  YYYY= -5.9890  ZZZZ= -5.1839  XXXY= 0
.0000
  XXXZ= 0.0000  YYYYX= 0.0000  YYYYZ= 0.0000  ZZZX= 0
.0000
  ZZZY= 0.0000  XXYY= -1.5810  XXZZ= -2.0167  YYZZ= -1
.9085
  XXYZ= 0.0000  YXXZ= 0.0000  ZZZY= 0.0000
N-N= 9.289187493137D+00 E-N=-1.990850536058D+02 KE= 7.585992283909D+01
Symmetry A' KE= 7.130832861150D+01
1280,2 94%

```

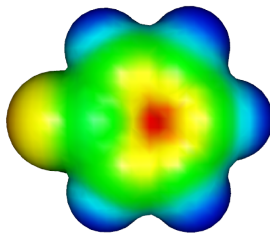
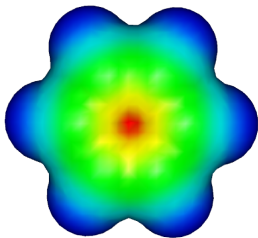
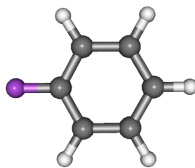
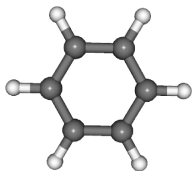
# Vizualizace

- Orbitaly



# Vizualizace

- Elektrostatický potenciál



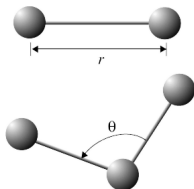
# Molekulová mechanika - formulace

- Popis molekul pomocí klasické mechaniky

# Molekulová mechanika - formulace

- Popis molekul pomocí klasické mechaniky
- Neuvažuje elektronovou strukturu
- Empirický potenciál ("forcefield")
- Obvykle se skádá z částí:
  - Vazebné členy - délky, úhly, torze
  - Elektrostatika
  - Repulze a disperze
- Coarse-grained modely

# Vazby, úhly



- Jednoduchý harmonický potenciál
- Parametry: rovnovážná hodnota, silová konstanta

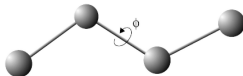
$$E_{bonds} = \sum_{ij} k_{ij} (r_{ij} - r_{ij}^0)^2$$

$$E_{ang} = \sum_{ijk} k_{ijk} (\theta_{ijk} - \theta_{ijk}^0)^2$$

- Nepopisuje disociaci vazeb
- Parametry tabelovány pro **atomové typy**



# Torzní úhly



- Periodické, různá četnost

$$E_{tor} = \sum_{ijkl} \sum_m k_{ijkl,m} (1 + \cos(n\phi - \phi^0))$$

- Někdy se používá i harmonický potenciál - dvojně vazby

# Elektrostatická interakce

- Mezi všemi atomy s výjimkou vázaných (až 1-3)

$$E_{ele} = \sum_i \sum_{j < i} \frac{1}{4\pi\epsilon} \frac{q_i q_j}{r_{ij}}$$

- Vyžaduje parciální atomové náboje
  - Lze spočítat z QM
  - Knihovny stavebních bloků - např. aminokyseliny v proteinech
- Zanedbává polarizaci

# Pauliho repulze, Londonova disperze

- Lennard-jonesův potenciál

$$E_{LJ} = \sum_i \sum_{j < i} \epsilon_{ij} \left[ \left( \frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left( \frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$$

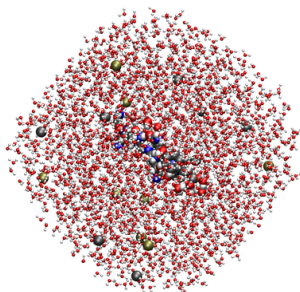
- Rovnovážná vzdálenost  $\sigma_{ij}$
- Energie minima  $\epsilon_{ij}$

# Molekulová mechanika - omezení

- Limitováno formou potenciálu
- Nepopisuje reakce
- Vyžaduje rozsáhlou parametrizaci
  - Specializované parametry - biomolekuly, materiály, ...
  - Univerzální parametry dávají jen přibližné výsledky

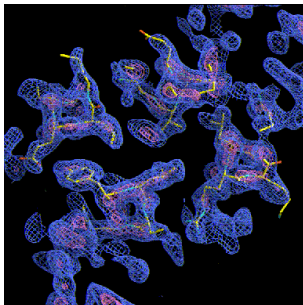
# Voda

- Většina chemie se odehrává v roztoku
- Explicitní
  - Jednotlivé molekuly, často přibližný popis
  - Periodické podmínky
- Dielektrické modely



# Použití

- Optimalizace struktur, frekvence, ...
- Molekulová dynamika
- Interpretace X-ray spekter
  - Měříme mapu elektronové hustoty
  - Optimalizujeme strukturu tak, aby odpovídala el. hustotě



# Molekulová dynamika

- Řešíme Newtonovy rovnice pro pohyb jader

$$x(t + \Delta t) = x(t) + v\Delta t + \frac{1}{2}a\Delta t^2$$

$$F = ma = -\frac{\partial E}{\partial x}$$

- Časový krok 1 fs
- Lze simulovat  $\sim$ ms (protein ve vodě)

# Použití

- Simulace dynamických dějů
- Analýza struktury
- Výpočet termodynamických veličin



# Výpočty volné energie

- Statický výpočet nestačí, nutný průměr přes soubor konfigurací
- V simulaci lze nastavit podmínky -  $T$ ,  $p$
- Nelze určit absolutní hodnoty  $G$ , umíme ale počítat  $\Delta G$