10.610 Quantum Chemistry Dry Lab

Li Ju

December 30, 1995

Contents

Chapter 1

Dry Lab 1: Hydrogenation of Carbon

From the *Handbook of Organic Chemistry* by John.A. Dean, we get the following structural data:

	$-C \equiv C-$:	**1.204 Å**	**Triple bond, simple**
Carbon-Carbon bond:	$-C = C-$:	**1.337 Å**	**Double bond, simple**
	$-C - C-$:	**1.541 Å**	**Single bond, simple**

Carbon-Hygrogen bond: C_2H_2: **1.059Å** $H - C - Y$: **1.096 Å** CH_4: **1.094 Å**

The bond length of H_2 is 0.7414 Å(1 a.u. = 0.5291771 Å). From *Thermochemical Data of Organic Compounds* by Pedley, we get the experimental enthalphy of formation for acetylene, ethylene and ethane:

$H_f(C_2H_2) = 228.2$ **kJ/mol** $H_f(C_2H_4) = 52.5$ **kJ/mol** $H_f(C_2H_6) = -83.8$ **kJ/mol**

From here we can calculate the heat of reaction for the following hydrogenation process:

```
          (Reaction 1): C2H2 + H2 = C2H4
So heat of reaction = 52.5-228.2 = -175.7 kJ/mol

          (Reaction 2): C2H4 + H2 = C2H6
So heat of reaction = -83.8-52.5 = -136.3 kJ/mol
```

We will calculate the above data using *ab initio* quantum chemistry package GAMESS (General Atomic and Molecular Electronic Structure Software), with increasing degrees of sophistication.

1.1 STO-3G Minimal Basis

STO-3G minimal basis means one STO (Slater-type orbital) for each occupied angular momentum state. 3G means every STO consists of three Guassians.

1) H_2

```
E(RHF)=  -1.1167143252 a.u.
Frequency =  4515.20 cm-1
COORDINATES OF SYMMETRY UNIQUE ATOMS (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0   0.0000000000   0.0000000000   0.3252537104
```

	E	H	G	CV	CP	S
	KJ/MOL	KJ/MOL	KJ/MOL	J/MOL-K	J/MOL-K	J/MOL-K
ELEC.	0.000	0.000	0.000	0.000	0.000	0.000
TRANS.	3.718	6.197	-28.831	12.472	20.786	117.487
ROT.	2.479	2.479	-1.120	8.314	8.314	12.072
VIB.	32.781	32.781	32.781	0.000	0.000	0.000
TOTAL	38.978	41.457	2.829	20.786	29.100	129.559

2) Acetylene: C_2H_2

```
E(RHF)=  -75.8562476987 a.u.
Frequencies (cm-1): 945.58      945.58      988.91      988.91
                    2497.35     3852.15     4023.64

COORDINATES OF ALL ATOMS ARE (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0   0.0000000000   0.0000000000  -1.6496165241
H           1.0   0.0000000000   0.0000000000   1.6496165241
C           6.0   0.0000000000   0.0000000000  -0.5842181757
C           6.0   0.0000000000   0.0000000000   0.5842181757

Carbon-Carbon bond length = 1.16843 ANGS
```

	E	H	G	CV	CP	S
	KJ/MOL	KJ/MOL	KJ/MOL	J/MOL-K	J/MOL-K	J/MOL-K
ELEC.	0.000	0.000	0.000	0.000	0.000	0.000
TRANS.	3.718	6.197	-38.342	12.472	20.786	149.386
ROT.	2.479	2.479	-10.985	8.314	8.314	45.160

```
 VIB.      85.628    85.628    85.093     6.954     6.954     1.793
 TOTAL     91.825    94.304    35.765    27.740    36.055   196.340
```

The *.inp* file for C_2H_2 look like this:

```
 $CONTRL SCFTYP=RHF RUNTYP=HESSIAN $END
!$CONTRL SCFTYP=RHF RUNTYP=OPTIMIZE NZVAR=5 $END
 $FORCE METHOD=ANALYTIC $END
!$BASIS  GBASIS=N31 NGAUSS=6 NDFUNC=1 NPFUNC=1 $END
 $BASIS  GBASIS=STO NGAUSS=3 $END

 $DATA
C2H2 (this is the title card)
DNH      4

 H           1.0   0.0000000000   0.0000000000   1.6498986849
 C           6.0   0.0000000000   0.0000000000   0.5930253261
 $END
 $ZMAT   IZMAT(1)=
        1,1,2,   1,1,3,  1,2,4, 5,1,2,4  5,2,1,3
 $END
 $LIBE  APTS(1) = 1.0,0.0,0.0,1.0,0.0,0.0 $END
```

Here $LIBE$ group is important. It specifies the two-fold degenerate *Linear Bend* plane of a 3-atom chain, which's specified in $ZMAT$ group behind 5: 5,1,2,4 5,2,1,3.

3) Ethylene: C_2H_4

```
 E(RHF) =  -77.0739547580 a.u.
 Frequencies(cm^-1): 990.36   1200.76  1205.94  1300.88  1461.21  1623.17
                     1752.63  2045.71  3568.93  3615.10  3735.98  3749.77
 Group: CNV 2
 COORDINATES OF SYMMETRY UNIQUE ATOMS (ANGS)
   ATOM   CHARGE       X              Y              Z
 ------------------------------------------------------------
 H           1.0   0.9157406662   0.0000000000   1.2291817802
 C           6.0   0.0000000000   0.0000000000   0.6529813867

 Carbon-Carbon bond length = 1.30596 ANGS

             E         H         G         CV        CP        S
          KJ/MOL    KJ/MOL    KJ/MOL   J/MOL-K   J/MOL-K   J/MOL-K
 ELEC.     0.000     0.000     0.000     0.000     0.000     0.000
 TRANS.    3.718     6.197   -38.620    12.472    20.786   150.317
```

```
ROT.       3.718     3.718   -16.005    12.472    12.472    66.152
VIB.     157.258   157.258   156.968     4.666     4.666     0.971
TOTAL    164.695   167.173   102.344    29.609    37.924   217.440
```

4) Ethane: C_2H_6

```
E(RHF) =  -78.3061796282 a.u.
Frequencies(cm^-1):   315.79   980.47   980.47   1193.30  1459.29   1459.29
                      1715.59  1762.43  1831.98  1831.98  1847.76   1847.76
      3569.72  3570.74  3746.30  3746.30  3758.48   3758.48

COORDINATES OF ALL ATOMS ARE (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0  -0.8796102553  -0.5078432177   1.1532386695
H           1.0   0.8796102553  -0.5078432177   1.1532386695
H           1.0   0.0000000000  -1.0156864354  -1.1532386695
H           1.0   0.8796102553   0.5078432177  -1.1532386695
H           1.0  -0.8796102553   0.5078432177  -1.1532386695
H           1.0   0.0000000000   1.0156864354   1.1532386695
C           6.0   0.0000000000   0.0000000000  -0.7688385534
C           6.0   0.0000000000   0.0000000000   0.7688385534

Carbon-Carbon bond length = 1.53768 ANGS


            E         H         G         CV        CP        S
         KJ/MOL    KJ/MOL    KJ/MOL   J/MOL-K   J/MOL-K   J/MOL-K
ELEC.     0.000     0.000     0.000     0.000     0.000     0.000
TRANS.    3.718     6.197   -38.878    12.472    20.786   151.183
ROT.      3.718     3.718   -16.589    12.472    12.472    68.113
VIB.    236.880   236.880   234.854    12.446    12.446     6.795
TOTAL   244.317   246.796   179.387    37.389    45.704   226.091
```

Now, we are going to calculate the heat of reaction from our quantum mechanical output. One important issue is that apart from the electronic contributions, the zero-point vibrational energies of the molecules is also important, that's why we calculated the Hessian at the equilibrium position (see *.inp* file for C_2H_2. The thermochemical data will be automatically printed out). Then

```
(Reaction 1): C2H2 + H2 = C2H4
Electronic contribution = -77.0739547580 - (-1.1167143252 - 75.8562476987) a.u.
                        = -0.10099273410000 a.u.
```

```
                        = -265.334 kJ/mol (1 a.u. = 2627.259 kJ/mol = 315900 K)

Others = 167.173 - (41.457 + 94.304) kJ/mol
       = 31.412 kJ/mol

So Hr = -265.334+31.412 = -233.922 kJ/mol

(Reaction 2): C2H4 + H2 = C2H6
Electronic contribution = -78.3061796282 - (-1.1167143252 - 77.0739547580) a.u.
                        =  -0.11551054500001 a.u.
                        =  -303.476 kJ/mol
Others = 246.796 - (41.457 + 167.173) kJ/mol
       = 38.166 kJ/mol
So Hr = -303.476+38.166 = -265.310 kJ/mol
```

1.2 $6-31G^{**}$ Basis Set

$6-31G^{**}$ means one STO (6 Gaussians) for core orbitals, two STO's (3G+1G) for valence electrons, three p polarization orbitals for H, Na, K etc., and six d polarization orbitals for other atoms.

1) H_2

```
E(RHF)=   -1.1313335881 a.u.
Bond length = 0.732572 Ang.
Frequency =  4635.10 cm^-1
```

	E	H	G	CV	CP	S
	KJ/MOL	KJ/MOL	KJ/MOL	J/MOL-K	J/MOL-K	J/MOL-K
ELEC.	0.000	0.000	0.000	0.000	0.000	0.000
TRANS.	3.718	6.197	-28.831	12.472	20.786	117.487
ROT.	2.479	2.479	-1.260	8.314	8.314	12.539
VIB.	27.724	27.724	27.724	0.000	0.000	0.000
TOTAL	33.921	36.400	-2.367	20.786	29.100	130.026

2) Acetylene: C_2H_2

```
E(RHF)=      -76.8218374417 a.u.
Frequencies (CM^-1): 799.29      799.29      877.08      877.08
                     2242.92     3585.59     3697.09
```

```
COORDINATES OF ALL ATOMS ARE (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0   0.0000000000   0.0000000000  -1.6498986849
H           1.0   0.0000000000   0.0000000000   1.6498986849
C           6.0   0.0000000000   0.0000000000  -0.5930253261
C           6.0   0.0000000000   0.0000000000   0.5930253261

Carbon-Carbon bond length = 1.18605 ANGS

            E         H         G         CV        CP        S
         KJ/MOL    KJ/MOL    KJ/MOL   J/MOL-K   J/MOL-K   J/MOL-K
ELEC.     0.000     0.000     0.000     0.000     0.000     0.000
TRANS.    3.718     6.197   -38.342    12.472    20.786   149.386
ROT.      2.479     2.479   -11.031     8.314     8.314    45.311
VIB.     77.752    77.752    76.851     9.927     9.927     3.021
TOTAL    83.949    86.428    27.478    30.713    39.027   197.719
```

3) Ethylene: C_2H_4

```
E(RHF) =  -78.0388414541 a.u.
Frequencies(cm^-1): 894.06   1092.29  1100.30  1151.41   1346.75  1488.06
    1597.78  1851.71  3297.82  3321.45   3375.45  3401.75

Group: CNV 2
COORDINATES OF SYMMETRY UNIQUE ATOMS (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0   0.9153898285   0.0000000000   1.2245415747
C           6.0   0.0000000000   0.0000000000   0.6582253212

Carbon-Carbon bond length = 1.31645 ANGS

            E         H         G         CV        CP        S
         KJ/MOL    KJ/MOL    KJ/MOL   J/MOL-K   J/MOL-K   J/MOL-K
ELEC.     0.000     0.000     0.000     0.000     0.000     0.000
TRANS.    3.718     6.197   -38.620    12.472    20.786   150.317
ROT.      3.718     3.718   -15.993    12.472    12.472    66.112
VIB.    143.447   143.447   142.992     6.662     6.662     1.528
TOTAL   150.884   153.363    88.379    31.605    39.919   217.958
```

4) Ethane: C_2H_6

```
E(RHF) =  -79.2382349981 a.u.
```

```
Frequencies(cm^-1): 327.27    884.77    884.77   1057.69   1328.88  1328.88
                    1531.38   1564.41   1627.36  1627.36   1632.41  1632.41
                    3178.64   3184.30   3232.37  3232.37   3257.39  3257.39

COORDINATES OF ALL ATOMS ARE (ANGS)
  ATOM   CHARGE       X              Y              Z
------------------------------------------------------------
H           1.0  -0.8765567132  -0.5060802544   1.1563233270
H           1.0   0.8765567132  -0.5060802544   1.1563233270
H           1.0   0.0000000000  -1.0121605087  -1.1563233270
H           1.0   0.8765567132   0.5060802544  -1.1563233270
H           1.0  -0.8765567132   0.5060802544  -1.1563233270
H           1.0   0.0000000000   1.0121605087   1.1563233270
C           6.0   0.0000000000   0.0000000000  -0.7633970447
C           6.0   0.0000000000   0.0000000000   0.7633970447

Carbon-Carbon bond length = 1.52679 ANGS
```

	E	H	G	CV	CP	S
	KJ/MOL	KJ/MOL	KJ/MOL	J/MOL-K	J/MOL-K	J/MOL-K
ELEC.	0.000	0.000	0.000	0.000	0.000	0.000
TRANS.	3.718	6.197	-38.878	12.472	20.786	151.183
ROT.	3.718	3.718	-16.563	12.472	12.472	68.025
VIB.	209.469	209.469	207.299	14.923	14.923	7.276
TOTAL	216.906	219.385	151.858	39.866	48.181	226.484

So,

```
(Reaction 1): C2H2 + H2 = C2H4
Electronic contribution = -78.0388414541 - (-1.1313335881 -76.8218374417) a.u.
                        = -0.08567042429999 a.u.
                        = -225.0784 kJ/mol
Others = 153.363 - (36.400 + 86.428) kJ/mol
       = 30.535 kJ/mol
So Hr = -225.0784 + 30.535 = -194.5434 kJ/mol


(Reaction 2): C2H4 + H2 = C2H6
Electronic contribution = -79.2382349981 - (-1.1313335881 -78.0388414541) a.u.
                        = -0.06805995590000 a.u.
                        = -178.8111 kJ/mol
Others = 219.385 - (36.400 + 153.363) kJ/mol
       = 29.6219 kJ/mol
So Hr = -178.8111 + 29.6219 = -149.189 kJ/mol
```

1.3 Perturbation Theory

The above calculations didn't take into account the effect of correlation, i.e., that the true ground state wavefunction of the system is not a pure single Slater-determinant. One way of treating the correlation is to use perturbation theory, which works well (and fast) when far from disassociation. Here we use Moller-Plesset 2nd-order perturbation theory, in the geometry minimized by $6-31G^{**}$ SCF, and it will imporve our heat of reaction. Here only the electronic contribution is considered. Zero-pt energy is relatively small and too time-consuming for MP2, so data from $6-31G^{**}$ is used.

1) H_2

```
 RESULTS OF MOLLER-PLESSET 2ND ORDER CORRECTION ARE
              E(0)=        -1.1313335881
              E(1)=         0.0
              E(2)=        -0.0263263739
            E(MP2)=        -1.1576599620
```

2) C_2H_2

```
 RESULTS OF MOLLER-PLESSET 2ND ORDER CORRECTION ARE
              E(0)=       -76.8218374417
              E(1)=         0.0
              E(2)=        -0.2577117239
            E(MP2)=       -77.0795491656
```

3) C_2H_4

```
 RESULTS OF MOLLER-PLESSET 2ND ORDER CORRECTION ARE
              E(0)=       -78.0388414541
              E(1)=         0.0
              E(2)=        -0.2779697309
            E(MP2)=       -78.3168111850
```

4) C_2H_6

```
 RESULTS OF MOLLER-PLESSET 2ND ORDER CORRECTION ARE
              E(0)=       -79.2382349981
              E(1)=         0.0
              E(2)=        -0.3051195685
            E(MP2)=       -79.5433545666
```

And

```
(Reaction 1): C2H2 + H2 = C2H4
Hr = -194.5434 + (-0.2779697309-(-0.0263263739-0.2577117239))*2627.259
   = -178.60 kJ/mol

(Reaction 2): C2H4 + H2 = C2H6
Hr = -149.189 + (-0.3051195685-(-0.0263263739-0.2779697309))*2627.259
   = -151.35 kJ/mol
```

1.4 Semi-Empirical Method: AM1

GAMESS incorporate MNDO,PM3,AM1 of the semi-empirical MOPAC package. The basic idea is to ignore most of the two-electron integrals in SCF and parametrize others from experimental data. To use it, just specify "MNDO", "PM3" or "AM1" in the $BASIS$ group.

1) H_2

```
Bong length = 0.67658244 ANGS
```

2) Acetylene: C_2H_2

```
HEAT OF FORMATION = 54.84891 KCAL/MOL = 229.6414 kJ/mol
Carbon-Carbon bond length = 1.19541 ANGS
```

3) Ethylene: C_2H_4

```
HEAT OF FORMATION = 16.47108 KCAL/MOL = 68.9611 kJ/mol
Carbon-Carbon bond length = 1.32594 ANGS
```

4) Ethane: C_2H_6

```
HEAT OF FORMATION = -17.41450 KCAL/MOL = -72.911 kJ/mol
Carbon-Carbon bond length = 1.50025 ANGS

    (Reaction 1): C2H2 + H2 = C2H4
Hr = 68.9611-229.6414 = -160.68 kJ/mol
```

```
    (Reaction 2): C2H4 + H2 = C2H6
Hr = -72.911-68.9611 =  -141.87 kJ/mol
```

The heat of formation result is good, but that must because all these common compounds are in the "training set" (parameter fitting) of the system. The geometry-prediction of semi-empirical methods is always very poor.

1.5 Density Functional Theory

Here I use **DMol**, a local density functional (LDA) quantum chemistry program, in the BIOSYM software package. Before executing **DMol**, we first construct the molecule using **Builder** module, a program whose output is a common object throughout BIOSYM. The graphical interface of BIOSYM is a great enhancement over GAMESS.

Constructing hydrocarbon compounds in **Builder** is especially easy. We use the Sketch:Edit in **Builder**, and by specifying the bond type to be "Single_Bond", "Double_Bond" or "Triple_Bond", we would get C_2H_6, C_2H_4 and C_2H_2 correspondingly. The hydrogen atoms would be automatically added by the program Sketch:3D_Convert according to the valence number, and in the correct geometry. The main function of Sketch:3D_Convert is to turn 2D object *Sketch* into 3D *Molecule*, the object which could be handled by **DMol**. After the conversion, we could check the geometry by Atom:List, and change element type or delete atoms by Atom:Replace, Delete, thus a general molecule could be formed. Then we could save our objects on disk by Molecule:Put, where two files *.car* and *.mdf* are generated.

We'll use (**DMol**:)Setup:Calculation:Optimize_Frequency to get our thermochemical data. The Parameters:Properties,Output items are in charge of property print-outs. The Parameters:Atomic_Basis_Set we use are DNP, i.e., double numerical plus polarization functions, equivalent to $6-31G^{**}$ in HF. The Parameters:Mesh_Points are chosen to be Medium. There can be further "non-local" gradient corrections in addition to LDA by changing items in Parameters:Functionals, however in this run it isn't used.

1) H_2

```
E(DFT,LDA)=   -1.142008330 a.u.
```

```
Bond length = 0.7675 Ang.
Frequency =  4635.10 cm^-1 (not available in {\bf DMol}, GAMESS used)
Zero point vibrational energy: 27.724 kJ/mol
```

2) Acetylene: C_2H_2

```
E(DFT,LDA) = -76.63775661 a.u.
Frequencies (CM^-1): 520.9      561.0      686.4
                     2001.4     3349.4     3430.3

Zero point vibrational energy: 15.0817 kcal/mol = 63.1441 kJ/mol

              ATOMIC  COORDINATES
            x          y          z

   C    -1.143224   0.003094   0.000003
   C     1.142855   0.012819   0.000009
   H    -3.184478  -0.002594  -0.000001
   H     3.184051   0.043970   0.010027

Carbon-Carbon bond length = 1.2098 ANGS
```

3) Ethylene: C_2H_4

```
E(DFT,LDA) =  -77.87547031 a.u.
Frequencies(cm^-1): 806.7    900.7   938.2   1028.1   1197.4   1269.4
    1514.2   1696.3  3008.3  3047.2   3096.1   3154.3

Zero point vibrational energy: 30.9607 kcal/mol = 129.6263 kJ/mol

              ATOMIC  COORDINATES
            x          y          z

   C     1.254870  -0.003052   0.039707
   C    -1.255945   0.001651   0.012393
   H     2.341037  -1.774452   0.029384
   H     2.347387   1.764123   0.072528
   H    -2.342095   1.773062   0.022327
   H    -2.348462  -1.765534  -0.010011

Carbon-Carbon bond length = 1.3288 ANGS
```

4) Ethane: C_2H_6

```
E(DFT,LDA) =  -79.09610285 a.u.
Frequencies(cm^-1): 416.2     721.3     797.5     947.8     1096.2    1230.2
                    1316.3    1381.2    1418.8    1433.2    1513.0    1570.3
                    2935.9    2946.9    2996.8    3010.1    3031.0    3088.7

Zero point vibrational energy: 44.9402 kcal/mol = 188.1556 kJ/mol

             ATOMIC  COORDINATES
           x          y          z

   C    -1.426661  -0.035557  -0.082267
   C     1.427226   0.034883   0.083783
   H    -2.091617  -1.616316  -1.274297
   H    -2.307881  -0.261440   1.797368
   H    -2.190852   1.720918  -0.916078
   H     2.193758  -1.723065   0.912027
   H     2.306628   0.264919  -1.796760
   H     2.092954   1.614980   1.286227

Carbon-Carbon bond length = 1.5132 ANGS
```

So,

```
(Reaction 1): C2H2 + H2 = C2H4
Electronic contribution = -77.87547031 - (-1.142008330 -76.63775661) a.u.
                        = -0.09570537 a.u.
                        = -251.44279 kJ/mol
Others = 129.6263-(27.724+63.1441) kJ/mol
       = 38.7581 kJ/mol
So Hr = -251.44279 + 38.7581 = -212.6845 kJ/mol


(Reaction 2): C2H4 + H2 = C2H6
Electronic contribution = -79.09610285 - (-1.142008330-77.87547031) a.u.
                        = -0.07862421 a.u.
                        = -206.56616 kJ/mol
Others = 188.1556 - (27.724+129.6263) kJ/mol
       = 30.8053 kJ/mol
So Hr = -175.76086  kJ/mol
```

It seems to me that LDA at this level is not as accurate as HF for these small molecules. However since it's an N^3 method, maybe it's better for larger calculations.

1.5.1 Doing More (Non-local DFT)

Since I'm not quite satisfied with the above calculations, especially with the electronic part, I decided to take some steps further:

- To include non-local interactions in Parameters:Functionals, with gradient corrected *energy* and *potential*. GC_Echange is set to B_88(Becke's 1988), GC_Correlation is set to LYP (Lee-Yang-Parr functional).
- In Parameters:Basis_and_Mesh, the Frozen_Orbitals is set to None (default is Inner_Core), Mesh_Points is set to Fine (default is Medium).
- In Optimize:Parameters:Tolerances, the Gradient is set to 0.0002, instead of the default value 0.001.
- Let the object generated by **Builder** be processed by **DMol**:Symmetry:Find_Pt_Group, with Snap_Symmetry option on, which will transform the molecule to its highest intended symmetry, so as to be the initial configuration for geometry optimization. (I don't know how to specify symmetry group for **DMol** runs. Doing this manually will ensure the final configuration to be in the correct symmetry, although it doesn't save time).

Since the zero-pt energy of the previous LDA run is consistent with that of $6-31G^{**}$ RHF, and is relatively small, we'll omit the calculation. Only the electronic part is significant.

1) H_2

```
E(DFT)=   -1.167561000 a.u.
Bond length = 0.7509 Ang.
```

2) Acetylene: C_2H_2

```
 E(DFT) = -77.323988977 a.u.

          Coordinates (Angstroms)
    ATOM       X           Y           Z
   1  C     0.000000    0.000000    0.606731
   2  C     0.000000    0.000000   -0.606731
```

```
   3  H     0.000000    0.000000    1.680594
   4  H     0.000000    0.000000   -1.680594
  Point Group: D*h  Number of degrees of freedom:   2

 Carbon-Carbon bond length = 1.2135 ANGS
```

3) Ethylene: C_2H_4

```
 E(DFT) =  -78.566135917 a.u.

          Coordinates (Angstroms)
    ATOM       X           Y           Z
   1  C     0.000000    0.669717    0.000000
   2  C     0.000000   -0.669717    0.000000
   3  H     0.000000    1.247174   -0.929247
   4  H     0.000000    1.247174    0.929247
   5  H     0.000000   -1.247174    0.929247
   6  H     0.000000   -1.247174   -0.929247
  Point Group: D2h  Number of degrees of freedom:   3

 Carbon-Carbon bond length = 1.3394 ANGS
```

4) Ethane: C_2H_6

```
 E(DFT) =  -79.789850252 a.u.

          Coordinates (Angstroms)
    ATOM       X           Y           Z
   1  C     0.000000    0.000000   -0.769459
   2  C     0.000000    0.000000    0.769459
   3  H    -0.888054   -0.512718   -1.171404
   4  H     0.000000    1.025437   -1.171404
   5  H     0.888054   -0.512718   -1.171404
   6  H    -0.888054    0.512718    1.171404
   7  H     0.000000   -1.025437    1.171404
   8  H     0.888054    0.512718    1.171404
  Point Group: D3d  Number of degrees of freedom:   3

 Carbon-Carbon bond length = 1.5389 ANGS
```

So,

```
(Reaction 1): C2H2 + H2 = C2H4
Electronic contribution = -78.566135917 - (-1.167561000-77.323988977) a.u.
                        = -0.074586 a.u.
                        = 195.956582 kJ/mol
Others = 38.7581 kJ/mol
So Hr =  195.956582 + 38.7581 = -157.198 kJ/mol


(Reaction 2): C2H4 + H2 = C2H6
Electronic contribution =  -79.789850252 - (-1.167561000-78.566135917) a.u.
                        = -0.056153 a.u.
                        = -147.529355 kJ/mol
Others  = 30.8053 kJ/mol
So Hr = -147.529355 + 30.8053 = -116.724  kJ/mol
```

Obviously this is a very good result (better still if the zero-pt vibrational energy is calculated directly, I guess). The geometric predictions are astonishingly accurate. For C_2H_6, the largest molecule, it took about 105.6 min. (without group theoretical simplification) on a Indy SGI workstation, certainly less than CI, maybe faster than MP2 optimizations. I'm satisfied with the current state of affairs.

Chapter 2

Dry Lab 4: Interatomic Interactions

2.1 Ne-Ne Interactions

We will use STO-3G, $6-31G^{**}$, $6-311G^{**}$ (Pople's "triple split", provided for H-Ne), MP2 and CI to calculate the potential energy surface of Ne-Ne. From *Reference Data on Atoms, Molecules and Ions* by A. A. Radzig and B. M. Smirnov (a very good book, by the way), page 408, it's found that the minimum of Ne-Ne interaction is at $3.1\dot{A}$, with well depth of 3.6 meV (41.8 K).

Now, for an isolated atom, the calculated energies are as follows:

RHF(STO-3G)	-126.6045250888 a.u.		
RHF(6-31G**)	-128.4744065199 a.u.	RHF(6-311G**)	-128.5226687496 a.u.
MP2(6-31G**)	-128.6247223475 a.u.	MP2(6-311G**)	-128.7392465733 a.u.
SDCI(6-31G**)	-128.587234364 a.u.	SDCI(6-311G**)	-128.618127326 a.u.
TCI (6-31G**)	-128.587866251 a.u.	TCI (6-311G**)	-128.619608202 a.u.
QCI (6-31G**)	-128.590683901 a.u.	QCI (6-311G**)	-128.621361426 a.u.

The reason I decided to use $6-311G^{**}$ instead of $6-31G^{**}$ is that CI calculation depends very much on the size of the basis we choose. For instance, I used $6-31G^{**}$ and did SDCI, but failed to observe binding. Also, QCI is more useful than TCI, as seen from the table, but it's so expensive on Ne-Ne that it is impossible for an Athena workstation. Thus, I use $6-311G^{**}$ and do TCI on Ne-Ne. $6-311G^{**}$ and its MP2 is also implemented for

comparison.

The *.inp* file of single-atom CI calculation is

```
 $CONTRL SCFTYP=CI $END
 $BASIS GBASIS=N311 NGAUSS=6  NDFUNC=1 NPFUNC=1 $END
 $GUESS GUESS= MOREAD NORB=19 $END
 $DRT  GROUP=C1  IEXCIT=3 NDOC=5 NVAL=5 NEXT=-1 $END
 $GUGDIA NSTATE = 2 $END
 $DATA
Ne
C1
NE          10.0       0.0000000000       0.0000000000       0.0000000000
 $END

 $VEC
  .
  .
  .
 $END
```

Here the $GUESS group is specified to be MOREAD, which means we read in orbitals generated by other calculations from the $VEC group, which's quoted from the *.dat* file of previous RHF calculation using $6-311G^{**}$. IEXCIT = 3 means we only allow single, double and triple excitations. On the other hand, FORS=.TRUE. means *fully optimized reaction space*, which will include all levels of excitations, but may not be complete (specifying SOCI=.TRUE. will ensure completeness of SDCI). It seems that FORS will take much more memory and machine time than IEXCIT=3.

NDOC indicates the number of doubly occupied orbitals of the *reference* configuration. Since neon has 10 electrons, all paired, NDOC=5. NVAL is the number of occupied orbitals in the reference configuration allowed to be excited, if NVAL=0, the software will interpretate it as *all* orbitals are allowed to be excited, thus here NVAL=5 will have the same effect. NEXT is the number of virtual orbitals of the reference configuration permitted to be excited into. NEXT=-1 means all are permitted.

NSTATE = 2 means we want 2 lowest solutions of the big CI matrix. The first one will of course be the ground state, the second one will be the lowest electronic excited state, etc. The output look like this

```
 STATE #    1  ENERGY =    -128.618127326

      CSF      COEF    OCCUPANCY (IGNORING CORE)
      ---      ----    --------- --------- -----
        1    0.990361  2222200000

 STATE #    2  ENERGY =    -126.956680627

      CSF      COEF    OCCUPANCY (IGNORING CORE)
      ---      ----    --------- --------- -----
       67    0.540237  2222100100
       69    0.634533  2212200100
      138    0.472768  2221201000
      233    0.259471  2221210000
```

Here CSF means various single determinants, denoted by a number. COEF are their coefficients in the eigenvector. Here in STATE #1 only CSF 1 is listed, because all others' coefficients are too small. The threshhold for listing could be changed by the program.

We now turn to calculating the interaction. The procedure could be automatically run by a shell script. Let's take the most complicated one, the CI calculation, as the example. First we construct the *.base* file:

```
 $CONTRL SCFTYP=CI $END
 $BASIS GBASIS=N311 NGAUSS=6  NDFUNC=1 NPFUNC=1 $END
 $GUESS GUESS= MOREAD NORB=38 $END
 $DRT  GROUP=C1  IEXCIT=3  NDOC=10 NVAL=8 NEXT=-1 $END
 $GUGDIA NSTATE = 2 $END
 $DATA
Ne-Ne interaction
C1
NE         10.0      0.0000000000      0.0000000000      0.0000000000
NE         10.0      0.0000000000      0.0000000000      variable_r
 $END
```

Then we use shell script "runCIlist",

```
#!/bin/csh
# usage: runlist base_file_name output_name <distances>

set base = $1
set out  = $2
```

```
set i=3
 while ( $i <=  $#argv )
   set  x=$argv[$i]

#  First run RHF 6-311G** to get the MO's
   rm nene6-311GSS.dat
   sed "s/variable_r/$x/" < nene6-311GSS.base > nene6-311GSS.inp
   rungms nene6-311GSS >& nene6-311GSS.log

   sed "s/variable_r/$x/" < $base.base > $base.inp
#  copy the MO's to $base.inp for CI calculation
   cat nene6-311GSS.dat | block ' $VEC' ' $END' >> $base.inp
   rm $base.dat
   rungms $base >& $base.log

   echo $x >> num
   grep " STATE #    1  ENERGY =" $base.log | cut -b 27-40 >> dat
   echo "Job R=$x done"
   @ i++
 end
paste num dat > $out.OUT
rm num dat nene6-311GSS.dat $base.dat core
```

Here we first run $6-311G^{**}$ RHF at a specified R to get the MO's, then we clip down the \$VEC group and combine it with the *.base* file to form an *.inp* file, and then let it run. In the end we pick out the energy output from the *.log* file, and add it to our $.OUT$ file.

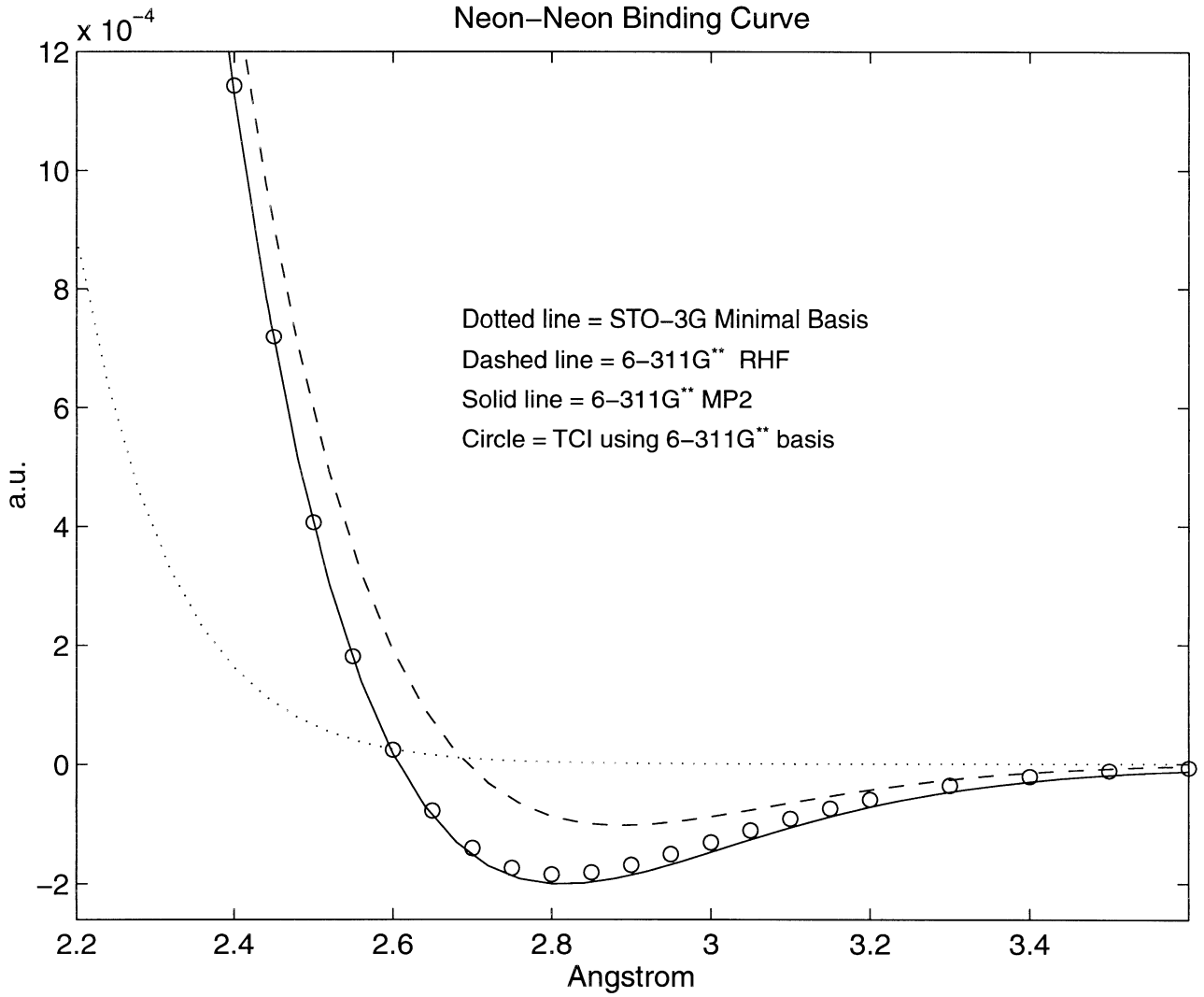
The results are shown in the figure of neon-neon binding curve. We see that the minimal STO-3G failed to show binding, while both MP2 and TCI on $6-311G^{**}$ do, and they agree extremely well. The position of the minimum is around 2.8 angstrom, the well depth is about $1.84E-4$ a.u. (58.2 K) for TCI, $2.00E-4$ a.u. (63.3 K) for MP2. The decay of interaction in long range is in the form of $1/R^6$, since if we take arbitary two points on the MP2 curve,

```
R = 4 A = 2.1167 a.u.      V = -4.2684e-06 a.u.
R = 8 A = 4.2334 a.u.      V = -6.1900e-08 a.u.
V(4)/V(8) =  68.9564
```

which's about $2^6 = 64$.

One can further explain the difference between experiment and calculations by observing that there is a fictitious binding effect in $6-311G^{**}$ RHF, due to a better basis set when atoms

come together. It can be seen that if we subtract part of RHF from MP2, the minimum position as well as the well depth might improve. However it is only a handwaving argument, the real progress should come after a larger basis set and higher excitation levels.



As our concluding remark, it is found that in order to get the correct binding behaviour for Ne-Ne (very weak), the size of the basis set is crucial. For instance, if we do MP2 on $6-31G^{**}$, we would get our minimum position at 2.6 Åand a well depth of $-6.628E-4$ a.u.(209.4K)! It almost made me think that my TCI calculation (which took me twice as much effort as all others combined) was inadequete. It was only after I found out the experimental value and did MP2 on $6-311G^{**}$ did I realized that it's a problem of basis set in my previous MP2 calculation.

2.2 Na-Cl (Ionic and Radical) Interactions

For this group we do UHF calculations, since each atom has odd number of electrons. From *Reference Data on Atoms, Molecules and Ions* by A. A. Radzig and B. M. Smirnov, page 381, the disassociation energy for NaCl was found to be 0.1543 a.u. (4.2 eV).

For isolated atom of each species, UHF with multiplicity of 2 is used:

```
Na: E(UHF,6-31G**) = -161.8414350882 a.u.
    E(MP2,6-31G**) = -161.8414350882 a.u.

Cl: E(UHF,6-31G**) = -459.4479639194 a.u.
    E(MP2,6-31G**) = -459.5524332776 a.u.
```

It's interesting to see whether NaCl is a singlet or triplet: after doing UHF calculations at several distances, we conclude that singlets have much lower energy:

	(UHF MULT=3)	(UHF MULT=1)
2.0 A	-621.0977276995	-621.3670160232
3.0 A	-621.1739363949	-621.3769375168
4.0 A	-621.1784689887	-621.3283837765
6.0 A	-621.1797300706	-621.2761367494

In doing UHF with MULT=1, we should specify MIX=.TRUE. in the $GUESS group, which means breaking the initial spatial orbital symmetry for up and down spins. However in this case I found that the final results of UHF MULT=1 are indentical to the RHF calculations at all distances. So for this case, the symmetrical solution is a true minimum.

Since an MP2 run will also give the unperturbed energy, we can write a shell script to collect both MP2 and $6-31G^{**}$ results in a single run:

```
#!/bin/csh
# usage: runlist base_file_name output_name <distances>

set base = $1
set out  = $2
set i=3
 while ( $i <=  $#argv )
```

```
   set  x=$argv[$i]
   sed "s/variable_r/$x/" < $base.base > $base.inp
   rungms $base >& $base.log
   echo $x >> num
   grep "E(MP2)= " $base.log | cut -b 27-50 >> dat

#  newly added. To unite GSS calculation and MP2
   grep "TOTAL ENERGY = " $base.log | cut -b 34-50 >> dat1

   echo "Job R=$x done"
   rm $base.dat
   @ i++
 end
paste num dat  > $out.OUT
paste num dat1 > GSS.OUT
rm num dat
```

The *.base* file is below:

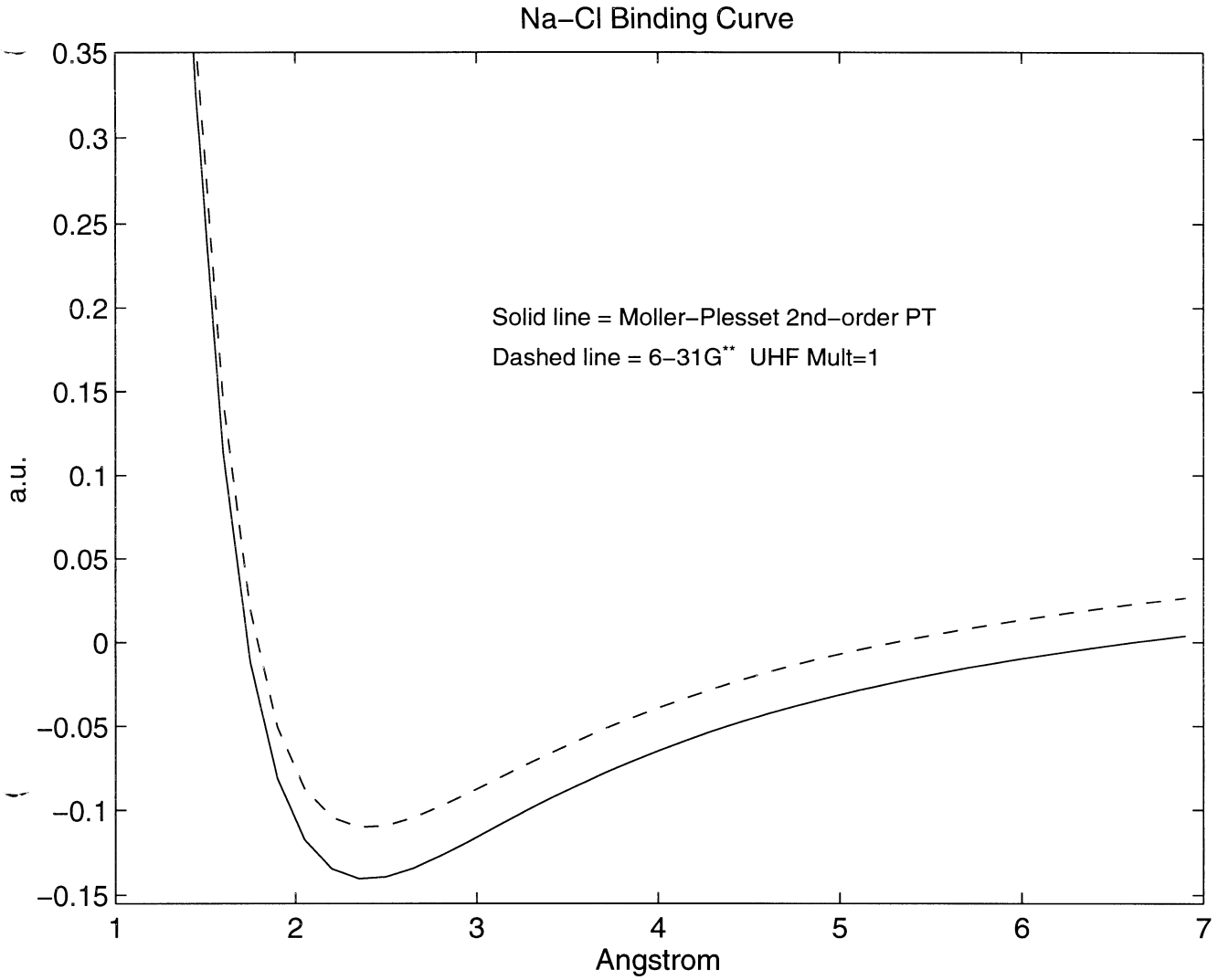
```
 $CONTRL MPLEVL=2 SCFTYP=UHF MULT=1 RUNTYP=ENERGY COORD=ZMT $END
 $BASIS  GBASIS=N31 NGAUSS=6  NDFUNC=1 NPFUNC=1 $END
 $GUESS  MIX=.T. $END
 $DATA
Na-Cl Interaction
CNV 4

 Na
 Cl 1 variable_r
 $END
```

The results are shown in the figure. One thing we can see is that although RHF is not that far away from MP2, it does not converge to zero when $R \to \infty$ as MP2 does, i.e., the energy of NaCl combined in UHF MULT=1 is greater than the sum of seperate Na and Cl atoms each with MULT=2. The reason for this is simple: like the case of H_2 disassociation, the correct wavefunction for NaCl at $R = \infty$ is

$$|\psi >= \frac{1}{\sqrt{2}}(|\phi_{Na}\bar{\phi}_{Cl} > +|\phi_{Cl}\bar{\phi}_{Na} >)$$

the combination of two Slater determinants (here bar means spin down), forming a singlet. Any method trying to describe it within a single determinant formalism is bound to fail. On the other hand, since MP2 takes into consideration the effect of correlation, it manages the situation quite well, as compared with experiment.


Na–Cl Binding Curve
Solid line = Moller–Plesset 2nd–order PT
Dashed line = 6–31G** UHF Mult=1
a.u.
Angstrom
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
−0.05
−0.1
−0.15
1
2
3
4
5
6
7