

Fortgeschrittenenpraktikum Physikalische Chemie SS07
Praktikumsversuch: Quantenchemie (Prof. Dr. Thorsten Klüner)

Ort: Nach Vereinbarung

Zeit: Nach Vereinbarung

Betreuer: Imed Mehdaoui (W3-0-032) und Matthias Mehring (W3-0-037)

Login: pcf

Passwort: Wird vom Assistenten mitgeteilt (pro Gruppe neues Password)

Literatur:

a) Praktikumsskript

b) Atkins Physikalische Chemie (Gruppentheorie)

c) F. Jensen, Introduction to Computational Chemistry

d) A. Szabo, N. Ostlund, Modern Quantum Chemistry

e) Einführende Literatur: Nobelvortrag John Pople

<http://nobelprize.org/chemistry/laureates/1998/pople-lecture.pdf>

Prüfungsrelevante Themen zum Antestat:

- Gruppentheorie

- Molekulare Schrödingergleichung

- Slaterdeterminanten

- Hartree-Fock-Theorie/SCF

- Konfigurationswechselwirkung

- Coupled Cluster

- Basissätze

- Z-Matrix

- Atomare Einheiten

- g03 Inputfiles (Aufbau, Optionen)

1.) Einführung in Unix

Unix-Kommandos und Tools „man, mkdir, cd, df, du, ssh, slogin, ls, pwd, ps, top, less, more, cat, tail, cp, mv, rm, chmod, find, diff, kill, hostname, grep, gnuplot, tar, gzip“

Aufgaben:

- Melden Sie sich mit Ihrem login und Ihrem password an dem Arbeitsplatzrechner an.
- Erstellen Sie ein Verzeichnis, in welchem der Praktikumsversuch durchgeführt wird (Namen der Teilnehmer und Datum sollten enthalten sein)
- Geben Sie den absoluten Pfad des Verzeichnisses an. Auf welchem Server befindet sich das von Ihnen angelegte Verzeichnis, wie viel Festplattenplatz steht auf dem Filesystem insgesamt zur Verfügung und wie viel Festplattenplatz könnten Sie nutzen (Angaben in GByte)?
- Melden Sie sich unter Verwendung einer secure shell (ssh) auf dem Server an, auf dem sich ihr Unterverzeichnis befindet und wechseln Sie in das von Ihnen auf der Workstation erstellten Unterverzeichnis.
- Wie viele Prozessoren und wie viel Hauptspeicher (in MByte) stehen Ihnen insgesamt zur Verfügung und wie groß ist die momentane Auslastung des Servers?
- Üben Sie den Umgang mit Unix, indem Sie einige der oben angegebenen Kommandos ausführen.

2.) Quantenchemie mit Gaussian03

Aufgaben: (alle in dem von Ihnen erstellten Unterverzeichnis durchzuführen)

2.1) Berechnung der Hartree-Fock-Energie und der Geometrie von Wasser (6-31G Basis):*

- Erstellen Sie z.B. mit dem Editor nedit ein Inputfile für das Programm g03 mit folgenden Zeilen und speichern Sie die Datei unter dem Namen H2O.com

```
-----  
#RHF/6-31G* POP=FULL GFInput IOP(6/7=3)
```

Meine erste ab initio Rechnung: HF-Energie von H2O

```
0 1  
O -0.464 0.177 0.0  
H -0.464 1.137 0.0  
H 0.441 -0.143 0.0  
-----
```

- Erklären Sie die Bedeutung der einzelnen Zeilen des Inputfiles.
- Starten Sie Ihre g03 Rechnung mit folgendem Kommando: g03<H2O.com>H2O.out
- Betrachten Sie das Outputfile (z.B. mit less oder nedit) und ermitteln Sie die Hartree-Fock-Gesamtenergie und die Partialladungen von Sauerstoff und Wasserstoff.

e) Führen Sie eine Geometrieoptimierung des Wassermoleküls durch. Erstellen Sie zu diesem Zweck ein neues Inputfile, das sich von Ihrem ersten Inputfile durch das zusätzliche Kommando OPT in der ersten Zeile unterscheidet. Ersetzen Sie die kartesischen Koordinaten durch eine genäherte Z-Matrix für das Wassermolekül und starten Sie danach das g03 Programm erneut, wobei Sie ein entsprechend dem Namen Ihres Inputfiles einen Namen Ihres Outputfiles wählen sollten. Je nach Wahl Ihrer Startgeometrie und Auslastung des Servers kann die Rechnung einige Minuten dauern.

f) Analysieren Sie das Outputfile und geben Sie die optimierte Geometrie und die dazugehörige Gesamtenergie an. Geben Sie die einzelnen Schritte der Geometrieoptimierung an, indem Sie die Gesamtenergie gegen die Schritte der Geometrieoptimierung auftragen. Dazu verwenden Sie das Programm gnuplot. Für das Protokoll können Sie natürlich auch Windowsprogramme wie Excel verwenden.

g) Führen Sie analoge Rechnungen mit einer STO-3G und einer 6-311+G** Basis durch. Was bedeuten diese Basissätze? Wie viele primitive bzw. kontrahierte Funktionen sind jeweils in diesen und im 6-31G*-Basissatz enthalten? Protokollieren und interpretieren Sie Ihre Ergebnisse (Gesamtenergie und optimierte Geometrie) im Sinne des Variationsprinzips und im Vergleich zum Experiment!

2.2) Protonenaffinität von NH_3 :

Ziel dieses Praktikumsversuchs ist die Berechnung der Protonenaffinität von Ammoniak in der Gasphase. Dies entspricht dem Gesamtenergieunterschied von NH_3 und NH_4^+ .

a) Welcher molekularen Punktgruppe sind diese Moleküle zugeordnet?

b) Konstruieren Sie die Z-Matrix dieser Moleküle mit Hilfe des Programms „molden“ (Einweisung durch den Assistenten).

c) Berechnen Sie auf Hartree-Fock-Niveau die Gesamtenergieunterschiede für die nachfolgenden Geometrien der Moleküle unter Verwendung eines STO-3G und eines 6-311G** Basissatzes.

NH_3 : $r(NH)=1.012 \text{ \AA}$, $\angle(HNH)=106,7^\circ$

NH_4^+ : $r(NH)=1.024 \text{ \AA}$, $\angle(HNH)=109.471^\circ$

d) Geben Sie die Protonenaffinität von Ammoniak in eV und kcal/mol an und vergleichen Sie Ihre Ergebnisse mit experimentellen Daten aus der Literatur (Diskussion!).

2.3) Berechnung einer Potentialkurve am Beispiel des N_2 -Moleküls:

Ziel dieses Praktikumsversuchs ist die Berechnung einer Potentialkurve für das Stickstoffmolekül.

a) Skizzieren Sie den Grundgedanken der Born-Oppenheimer-Näherung *quantitativ* ausgehend von der molekularen Schrödingergleichung und diskutieren Sie die Gültigkeit dieser Näherung! Informieren Sie sich dazu in Lehrbüchern oder in der Literatur.

b) Berechnen Sie die Potentialkurve des N_2 -Moleküls auf RHF/cc-pVTZ Niveau. Berechnen Sie dazu die Gesamtenergie als Funktion des Abstands der beiden

Stickstoffatome (von 0.7 Å bis 5.0 Å), wobei die Schrittweite 0.1 Å betragen soll. Sie können diese Rechnungen durch die Verwendung eines shell-skripts oder durch Schreiben eines FORTRAN-Programms automatisieren. Informationen erhalten Sie ggf. beim Assistenten.

c) Tragen Sie die Energie des N₂-Moleküls als Funktion des Abstands der beiden Stickstoffatome auf (z.B. mit gnuplot). Ermitteln Sie den Gleichgewichtsabstand des Moleküls.

d) Optimieren Sie nun die Geometrie des N₂-Moleküls automatisch durch Verwendung des OPT-Kommandos im g03 Inputfile und vergleichen Sie die Ergebnisse mit den Resultaten aus 2.2c). Diskutieren Sie die Ergebnisse im Vergleich zum Experiment und begründen Sie die möglicherweise auftretenden Abweichungen hinsichtlich der verwendeten ab initio Theorie (RHF).

e) Erstellen sie ein MO-Diagramm des N₂-Moleküls beim Gleichgewichtsabstand, indem Sie die Orbitalenergien (anzugeben in eV) der kanonischen besetzten und der niedrigsten unbesetzten RHF-Orbitale relativ zueinander auftragen. Diskutieren Sie Ihr Ergebnis im Vergleich zur Literatur (z.B. Atkins „Physikalische Chemie“ oder Lehrbücher der anorganischen Chemie) Visualisieren Sie die Orbitale unter Verwendung des Programms molden. Legen Sie Ausdrücke der Orbitale Ihrem Protokoll bei.

f) Optimieren Sie die Geometrie des N₂-Moleküls automatisch durch Verwendung des OPT-Kommandos im g03 Inputfile auf RHF, CISD, CCSD und CCSD(T) Niveau mit einer cc-pVTZ-Basis und vergleichen Sie die Ergebnisse mit den Resultaten aus 2.2d) bzw. dem Experiment. Wieviele primitive und kontrahierte Funktionen hat diese Basis für jeden Drehimpuls? **Achtung!** Diese Rechnungen sind ressourcenintensiv, daher wählen Sie eine sinnvolle Startgeometrie und sprechen Sie Ihr Vorgehen mit dem Assistenten ab. Erhöhen Sie die Hauptspeicheranforderung auf 500 MB durch folgendes Kommando, das in der ersten Zeile Ihrer input-files stehen muss:

```
%MEM=800MB
```

Rechnen Sie „im Hintergrund“, indem Sie Ihre Rechnungen wie folgt starten:

```
g03<name_ihres_input_files>name_ihres_output_files &
```

Kontrollieren Sie mit „top“ und „tail -f name_ihres_output_files“ den Fortschritt der Rechnungen. Wieviel Rechenzeit haben Sie am Ende verbraucht?

2.4) Geometrieoptimierung organischer Moleküle

Ziel dieses Praktikumsversuches ist die Geometrieoptimierung und die Berechnung von Schwingungsfrequenzen einfacher organischer Moleküle auf RHF/6-31G* Niveau.

a) Führen Sie eine Geometrieoptimierung des Ethen Moleküls (C₂H₄) durch. Geben Sie die optimierte Geometrie, die Gesamtenergie und die Schwingungsfrequenzen an. Zur Berechnung der Frequenzen verwenden Sie das Kommando **FREQ** in der ersten Zeile des g03 Inputfiles. Verwenden Sie zur Analyse der Ergebnisse das Programm „molden“.

b) Führen Sie eine zu Aufgabe 2.4a) analoge Rechnung von Fluorethen (C_2H_3F) durch. Geben Sie die optimierte Geometrie, die Gesamtenergie und die Schwingungsfrequenzen an und interpretieren Sie Ihre Ergebnisse im Vergleich zu den Resultaten aus Aufgabe 2.4a). Verwenden Sie zur Analyse der Ergebnisse das Programm „molden“.

2.5) *Size-Consistency*:

a) Berechnen Sie sowohl mit Hartree-Fock als auch mit Konfigurationswechselwirkung mit Einfach- und Doppelanregungen (CISD) die Gesamtenergie eines Systems aus zwei Wasserstoffmolekülen (interner H-H-Abstand 1.401 Bohr, Basis 6-31-G**) bei einem von Ihnen gewählten Abstand, bei dem die beiden Moleküle nicht wechselwirken.

b) Vergleichen Sie die von Ihnen berechneten Energien mit denen der Monomere und diskutieren Sie das Ergebnis.

3.) Sicherung der Ergebnisse:

a) Erstellen Sie ein tar-file der von Ihnen erzeugten Dateien und komprimieren Sie es mit Hilfe des Programms gzip.

b) Schicken Sie sich eine email, die das tarfile als Anhang enthält, so dass Sie ggf. zu Hause auf Ihre Files zugreifen können.

Hinweise zum
Fortgeschrittenenpraktikum Physikalische Chemie SS07
Praktikumsversuch: Quantenchemie (Prof. Dr. Thorsten Klüner)

1.) Einführung in Unix

Hilfe zu den einzelnen Unix-Kommandos erhalten sie mit: **man *Befehl***

Möchten Sie sich beispielsweise die Hilfe zum Befehl **df** anzeigen lassen, dann geben Sie in eine Shell ein: **man df**

a)

Login name: pcf
Passwort: „Wird vom Assistenten ausgegeben“

b)

mkdir - Verzeichnisse anlegen.
mkdir name1_name2_Datum; legt ein Verzeichnis mit dem Namen name1_name2_Datum an.

c)

pwd - zeigt den absoluten Pfad an.

df - zeigt die Belegung der gemounteten Filesysteme an.
df -k; zeigt die Belegung der gemounteten Filesysteme in Kilobytes an.

d)

ssh - Login auf einem Entfernten Rechner.
ssh name_des_rechners; einloggen auf dem Rechner name_des_rechners.

cd - in ein Verzeichnis wechseln.
cd name1_name2_Datum; wechselt in das Verzeichnis name1_name2_Datum.

e)

top - Verbrauchs-Informationen von Prozessen ausgeben.

f)

Üben Sie den Umgang mit Unix, indem Sie die unten aufgeführten Befehle der Reihe nach ausprobieren.

Öffnen Sie zunächst den Texteditor **nedit** und schreiben Sie 2-3 Zeilen beliebigen Text. In die letzte Zeile schreiben Sie „SCF Done“. Speichern Sie dann unter dem Namen Datei1 ab.

touch - legt eine (leere) Datei an bzw. ändert den Zeitstempel einer Datei.
touch Datei2; besteht Datei2 noch nicht, wird diese neu angelegt.

ls - Verzeichnis-Inhalte auflisten.

less - Dateien anzeigen.
less *Datei1*; zeigt *Datei1* an.

cp - Dateien und Verzeichnisse kopieren.
cp *Datei1 Datei2*; legt eine Kopie von *Datei1* unter dem Namen *Datei2* ab.

mv - Dateien und Verzeichnisse umbenennen oder verschieben.
mv *Datei2 Datei3*; benennt *Datei2* in *Datei3* um.

tar - Dateien und Verzeichnisse archivieren.
tar -cf *Datei1.tar Datei1*; erstellt das Archiv *Datei1.tar* aus *Datei1*.
tar -xf *Datei1.tar*; extrahiert das Archiv *Datei1.tar*.

gzip - Dateien komprimieren.
gzip *Datei1.tar*; erzeugt die komprimierte Datei *Datei1.tar.gz* aus *Datei1.tar*.

gunzip - Dateien dekomprimieren.
gunzip *Datei1.tar.gz*; dekomprimiert *Datei1.tar.gz*.

ps – Prozessinformationen anzeigen lassen.
ps -ef; zeigt Prozessinformationen ausführlich an.

kill - einen Prozess beenden.
kill *123456789*; beendet Prozess mit der PID *123456789*.

top - Verbrauchs-Informationen von Prozessen ausgeben.
top; zeigt mehr Informationen als **ps -ef** an (Verbrauch von Memory, CPU-Zeit).

grep - Zeilen aus Dateien oder einem Dateistrom herausfiltern.
grep '*SCF Done*' *Datei1*; Filtert den String '*SCF Done*' aus *Datei1* heraus.
grep '*SCF Done*' *Datei1* > *out* ; Filtert den String '*SCF Done*' aus *Datei1* heraus und leitet das Ergebnis in die Datei *out* um.

diff - Dateien miteinander vergleichen.
diff *Datei1 Datei3*; vergleicht *Datei1* und *Datei3*.

rm – Dateien und Verzeichnisse löschen.
rm *Datei3*; löscht *Datei3*.

2.) Quantenchemie mit Gaussian03

2.1) a)

Hinweis: Fügen Sie am Ende der Datei zwei Leerzeilen ein, da das Programm g03 sonst mit einer Fehlermeldung abbricht.

f)

Um ein Bild im postscript-Format mit **gnuplot** zu erstellen gehen Sie folgendermaßen vor:

- Filtern Sie zunächst die Hartree-Fock Energien der einzelnen Schritte mit dem Befehl **grep** heraus und leiten diese in die Datei *H2O_opt.rhf* um:

```
grep 'SCF Done' name_der_datei.out > H2O_opt.rhf
```

- Schauen Sie sich die Datei mit **less** an. In der 5. Spalte (jede Spalte wird durch Leerzeichen getrennt) sollten die Hartree-Fock Energien stehen.
- Führen Sie den Befehl **gnuplot** aus.
- Um die Gesamtenergie gegen die Schritte der Geometrieoptimierung aufzutragen führen Sie dann folgenden Befehl aus:
plot 'H2O_opt.rhf' using 5 with linespoints pointsize 2
- Formatieren Sie das Bild mit beispielsweise folgenden Befehlen:
set xlabel 'Anzahl der Schritte'
set ylabel 'Energie / Hartree'
set xtics 0,1,10
set title 'Bild zu Aufgabe 2.1 f)'
- Mit **replot** können Sie sich das Ergebnis der Befehle anschauen.
- Um das Bild im postscript-Format zu speichern führen Sie anschließend folgende Befehle der Reihe nach aus:
set terminal postscript
set output 'H2O_opt.rhf.ps'
replot
set term x11
- Beenden Sie dann das Programm mit **quit**.
- Betrachten Sie das Ergebnis mit **gsview**:
gsview H2O_opt.rhf.ps
- Ausführliche Informationen zu **gnuplot** finden Sie auf der **gnuplot** Homepage:
<http://www.gnuplot.info/>.
- Einen Grundkurs für **gnuplot** mit vielen nützlichen Tipps finden Sie beispielsweise hier:
<http://userpage.fu-berlin.de/~voelker/gnuplotkurs/gnuplotkurs.html>

2.2) b)

- Starten Sie das Programm **molden** mit dem Befehl: **molden4.3.irix &** (bzw. **molden4.0.linux**, wenn Sie auf einem Linux-Rechner arbeiten); Es öffnet sich das **Molden Control** Fenster (blau) und das **MOLDEN** Fenster (schwarz).

Molden Control

- Klicken Sie im **Molden Control** Fenster zunächst auf den Button **Solid** und wählen Sie **Ball and Stick** aus. Anschließend klicken Sie auf den Button **Shade**.
- Um den Z-Matrix Editor zu starten klicken Sie auf den Button **ZMAT Editor**.

Zmatrix editor

- Im **Zmatrix editor** klicken Sie auf den Button **Add Line** und wählen Sie ein Stickstoff-Atom aus. Im **MOLDEN** Fenster erscheint ein Stickstoff-Atom.

- Klicken Sie erneut auf den Button **Add Line**, wählen Sie ein Wasserstoff-Atom aus und klicken Sie dann auf das Stickstoff-Atom im **MOLDEN** Fenster.
- Klicken Sie nochmals auf den Button **Add Line** und wählen Sie ein Wasserstoff-Atom aus. Anschließend klicken Sie zuerst auf das Stickstoff-Atom und erst dann auf das Wasserstoff-Atom.
- Verfahren Sie analog für das dritte Wasserstoff-Atom. Achten Sie auf die Hinweise, die Ihnen eventuell das Programm gibt.
- Passen Sie die Winkel, falls erforderlich, im **Zmatrix Editor** an, um eine vernünftige Startgeometrie zu erhalten.
- Um eine Datei mit der Z-Matrix auszugeben, klicken Sie im **Zmatrix Editor** zunächst auf den Button **Gaussian**, geben dann einen Dateinamen in das Feld **File name** ein und klicken anschließend auf den Button **Write Z-Matrix**.
- Ausführliche Informationen zum Programm erhalten Sie auf der **molden** Homepage: <http://www.cmbi.ru.nl/molden/molden.html>

2.3) b)

Für diese Aufgabe können Sie sich eine Vorlage eines Shell-Skriptes kopieren, das sich im Verzeichnis `/h2/home/pcf/SKRIPTE` befindet und den Dateinamen **N2_ccpvtz.tcsh_template** hat. Kopieren Sie sich das Shell-Skript zunächst in Ihr Verzeichnis, bearbeiten Sie es dann entsprechend Ihren Anforderungen und führen es dann mit `./N2_ccpvtz.tcsh_template` aus. Falls Sie die Fehlermeldung **Permission denied** erhalten, müssen Sie die Ausführungsrechte der Datei mit folgendem Befehl ändern: **chmod u+x N2_ccpvtz.tcsh_template**.

c)

Für diese Aufgabe können Sie wieder die Programme **grep** und **gnuplot** verwenden.

Um mehrere Dateien gleichzeitig zu filtern verwenden Sie einen Stern (*) im Dateinamen nach folgendem Schema:

```
grep 'SCF Done' N2_DN2*out > N2_Potentialkurve.rhf
```

Editieren Sie dann die Datei `N2_Potentialkurve.rhf`, so dass die Bindungsabstände des N_2 -Moleküls immer in einer eigenen Spalte stehen (wichtig zum Plotten mit **gnuplot**).

Vorher: `N2_DN2_2.5_ccpvtz.out: SCF Done: E(RHF) = -108.148474879 A.U. after 5 cycles`
 Nachher: `N2_DN2 2.5 ccpvtz.out: SCF Done: E(RHF) = -108.148474879 A.U. after 5 cycles`

Konsultieren Sie Punkt 2.1 f) für die Verwendung von **gnuplot** (Hinweis: Um Spalte 2 und Spalte 8 einer Datei zu benutzen verwenden Sie den Befehl **using 2:8**; so können Sie x- und y-Achse definieren).

e)

Starten Sie das Programm **molden** mit dem Befehl: **molden4.3.irix &**

Lesen Sie die entsprechende Datei ein, indem Sie im **Molden Control** Fenster zunächst auf den Button **Read** klicken und dann die entsprechende Datei auswählen.

Um Orbitale zu visualisieren, klicken Sie zuerst auf den Button **Dens. Mode**, anschließend auf den Button **Space** und geben für **Contour Value** 0.3 ein. Probieren Sie auch andere Werte für **Contour Value** aus.

Die verschiedenen Orbitale lassen sich auswählen, wenn Sie im **Molden Control** Fenster auf den Button **Orbital** klicken.

Bilder der Orbitale lassen sich im postscript-Format ausgegeben, indem Sie im **Molden Control** Fenster auf den Button **postscript** klicken.

2.4) a) + b) + c)

Verwenden Sie das Programm **molden**, um sich die IR- und Raman-Spektren anzeigen zu lassen (Button **Norm. Mode**), geben Sie die Spektren im postscript-Format aus und fügen Sie diese Ihrem Protokoll bei.

Zur Analyse der Spektren ist es sehr hilfreich, wenn Sie sich mit **molden** die Normalschwingungen visualisieren. So lassen sich die Peaks in den Spektren sehr leicht zuordnen.

2.5)

Verwenden Sie für diese Aufgabe den 6-31G** Basissatz.

Die Einheit Bohr können Sie mit der Option **Units=AU** in der ersten Zeile Ihres Input-Files verwenden.

3.)

Bevor Sie die Ergebnisse sichern, kontrollieren Sie alle Outputs nach Fehlermeldungen. Dies geht sehr einfach mit dem grep Befehl: **grep -i 'error' *.out**

Eine beliebte Fehlerquelle ist auch die Verwendung falscher Geometrien. Dies lässt sich häufig vermeiden, wenn Sie im Output die Punktgruppe des Moleküls überprüfen, die das Programm g03 automatisch bestimmt. Für C₂H₄ sollte g03 die Punktgruppe D_{2h} bestimmt haben. Ist dies nicht der Fall, so haben Sie vermutlich die Geometrie falsch definiert.

4.)

Fügen Sie die Inputs dem Protokoll als Anhang bei.

5 Basis Sets

Ab initio methods try to derive information by solving the Schrödinger equation without fitting parameters to experimental data. Actually, *ab initio* methods also make use of experimental data, but in a somewhat more subtle fashion. Many different approximate methods exist for solving the Schrödinger equation, and the one to use for a specific problem is usually chosen by comparing the performance against known experimental data. Experimental data thus guides the selection of the computational model, rather than directly entering the computational procedure.

One of the approximations inherent in essentially all *ab initio* methods is the introduction of a basis set. Expanding an unknown function, such as a molecular orbital, in a set of known functions is not an approximation, if the basis is complete. However, a complete basis means that an infinite number of functions must be used, which is impossible in actual calculations. An unknown MO can be thought of as a function in the infinite coordinate system spanned by the complete basis set. When a finite basis is used, only the components of the MO along those coordinate axes corresponding to the selected basis can be represented. The smaller the basis, the poorer the representation. The type of basis functions used also influence the accuracy. The better a single basis function is able to reproduce the unknown function, the fewer are basis functions necessary for achieving a given level of accuracy. Knowing that the computational effort of *ab initio* methods scales formally as at least M^4 , it is of course of prime importance to make the basis set as small as possible without compromising the accuracy.¹

5.1 Slater and Gaussian Type Orbitals

There are two types of basis functions (also called *Atomic Orbitals*, AO, although in general they are not solutions to an atomic Schrödinger equation) commonly used in electronic structure calculations: *Slater Type Orbitals* (STO) and *Gaussian Type Orbitals* (GTO). Slater type orbitals² have the functional form

$$\chi_{\zeta,n,l,m}(r, \theta, \varphi) = N Y_{l,m}(\theta, \varphi) r^{n-1} e^{-\zeta r} \quad (5.1)$$

N is a normalization constant and $Y_{l,m}$ are the usual spherical harmonic functions. The exponential dependence on the distance between the nucleus and the electron mirrors the exact orbitals for the hydrogen atom. However, STOs do not have any radial nodes.

nodes in the radial part are introduced by making linear combinations of STOs. The exponential dependence ensures a fairly rapid convergence with increasing number of functions, however, as noted in section 3.5, the calculation of three- and four-centre two-electron integrals cannot be performed analytically. STOs are primarily used for atomic and diatomic systems where high accuracy is required, and in semi-empirical methods where all three- and four-centre integrals are neglected.

Gaussian type orbitals³ can be written in terms of polar or cartesian coordinates

$$\begin{aligned} \chi_{\zeta,n,l,m}(r, \theta, \varphi) &= N Y_{l,m}(\theta, \varphi) r^{(2n-2-l)} e^{-\zeta r^2} \\ \chi_{\zeta,l_x,l_y,l_z}(x, y, z) &= N x^{l_x} y^{l_y} z^{l_z} e^{-\zeta r^2} \end{aligned} \quad (5.2)$$

where the sum of l_x, l_y and l_z determines the type of orbital (for example $l_x + l_y + l_z = 1$ is a p-orbital). Although a GTO appears similar in the two sets of coordinates, there is a subtle difference. A d-type GTO written in terms of the spherical functions has five components ($Y_{2,2}, Y_{2,1}, Y_{2,0}, Y_{2,-1}, Y_{2,-2}$), but there appear to be six components in the Cartesian coordinates ($x^2, y^2, z^2, xy, xz, yz$). The latter six functions, however, may be transformed to the five spherical d-functions and one additional s-function ($x^2 + y^2 + z^2$). Similarly, there are 10 Cartesian "f-functions" which may be transformed into seven spherical f-functions and one set of spherical p-functions. Modern programs for evaluating two-electron integrals are geared to Cartesian coordinates, and they generate pure spherical d-functions by transforming the six Cartesian components to the five spherical functions. When only one d-function is present per atom the saving by removing the extra s-function is small, but if many d-functions and/or higher angular moment functions (f-, g-, h- etc. functions) are present, the saving can be substantial. Furthermore, the use of only the spherical components reduces the problems of linear dependence for large basis sets, as discussed below.

The r^2 dependence in the exponential makes the GTO inferior to the STOs in two aspects. At the nucleus the GTO has zero slope, in contrast to the STO which has a "cusp" (discontinuous derivative), and GTOs have problems representing the proper behaviour near the nucleus. The other problem is that the GTO falls off too rapidly far from the nucleus compared with an STO, and the "tail" of the wave function is consequently represented poorly. Both STOs and GTOs can be chosen to form a complete basis, but the above considerations indicate that more GTOs are necessary for achieving a certain accuracy compared with STOs. A rough guideline says that three times as many GTOs as STOs are required for reaching a given level of accuracy. The increase in number of basis functions, however, is more than compensated for by the ease by which the required integrals can be calculated. In terms of computational efficiency, GTOs are therefore preferred, and used almost universally as basis functions in electronic structure calculations. Furthermore, essentially all applications take the GTOs to be centred at the nuclei. For certain types of calculation the centre of a basis function may be taken not to coincide with a nucleus, for example being placed at the centre of a bond.

5.2 Classification of Basis Sets

Having decided on the type of function (STO/GTO) and the location (nuclei), the most important factor is the number of functions to be used. The smallest number of functions

possible is a *minimum basis set*. Only enough functions are employed to contain all the electrons of the neutral atom(s). For hydrogen (and helium) this means a single s-function. For the first row in the periodic table it means two s-functions (1s and 2s) and one set of p-functions ($2p_x$, $2p_y$ and $2p_z$). Lithium and beryllium formally only require two s-functions, but a set of p-functions is usually also added. For the second row elements, three s-functions (1s, 2s and 3s) and two sets of p-functions (2p and 3p) are used.

The next improvement in the basis sets is a doubling of all basis functions, producing a *Double Zeta (DZ)* type basis. The term zeta stems from the fact that the exponent of STO basis functions is often denoted by the greek letter ζ . A DZ basis thus employs two s-functions for hydrogen (1s and 1s'), four s-functions (1s, 1s', 2s and 2s') and two p-functions ($2p$ and $2p'$) for first row elements, and six s-functions and four p-functions for second row elements. The importance of a DZ over a minimum basis can be illustrated by considering the bonding in HCN (Figure 5.1). The C-H bond will primarily consist of the hydrogen s-orbital and the p_z -orbital on C. The π -bond between C and N will consist of the p_x (and p_y) orbitals of C and N. The π -bond will have a more diffuse electron distribution than the C-H σ -bond. The optimum exponent for the carbon p-orbital will thus be smaller for the x -direction than for the z -direction. If only a single set of p-orbitals is available (minimum basis) a compromise will be necessary. A DZ basis, however, has two sets of p-orbitals with different exponents. The tighter function (larger exponent) can enter the C-H σ -bond with a large coefficient, while the more diffuse function (small exponent) can be used primarily for describing the C-N π -bond. Doubling the number of basis functions thus allows for a much better description of the fact that the electron distribution is different in different directions.

The chemical bonding occurs between valence orbitals. Doubling the 1s-functions in for example carbon allows for a better description of the 1s-electrons. However, the 1s-orbital is essentially independent of the chemical environment, being very close to the atomic case. A variation of the DZ type basis only doubles the number of valence orbitals, producing a *split valence basis*. In actual calculations a doubling of the core orbitals would rarely be considered, and the term DZ basis is also used for split valence basis sets (or sometimes VDZ, for valence double zeta).

The next step up in basis set size is a *Triple Zeta (TZ)*. Such a basis contains three times as many functions as the minimum basis, i.e. six s-functions and three p-functions for the first row elements. Some of the core orbitals may again be saved by only splitting the valence, producing a *triple split valence* basis set. Again the term TZ is used to cover both cases. The names *Quadruple Zeta (QZ)* and *Quintuple Zeta (5Z, not QZ)* for the next levels of basis sets are also used, but large sets are often given explicitly in terms of the number of basis functions of each type.

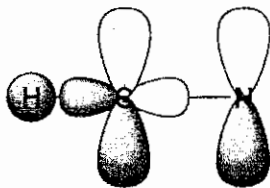


Figure 5.1 A double zeta basis allows for different bonding in different directions

So far only the number of s- and p-functions for each atom (first or second row in the periodic table) has been discussed. In most cases higher angular momentum functions are also important, these are denoted *polarization functions*. Consider again the bonding in HCN in Figure 5.1. The C-H bond is primarily described by the hydrogen s-orbital(s) and the carbon s- and p_z -orbitals. It is clear that the electron distribution along the bond will be different than that perpendicular to the bond. If only s-functions are present on the hydrogen, this cannot be described. However, if a set of p-orbitals is added to the hydrogen, the p_z component can be used for improving the description of the H-C bond. The p-orbital introduces a polarization of the s-orbital(s). Similarly, d-orbitals can be used for polarizing p-orbitals, f-orbitals for polarizing d-orbitals etc. Once a p-orbital has been added to a hydrogen s-orbital, it may be argued that the p-orbital now should be polarized by adding a d-orbital, which should be polarized by an f-orbital, etc. For single determinant wave functions, where electron correlation is not considered, the first set of polarization functions (i.e. p-functions for hydrogen and d-functions for heavy atoms) is by far the most important, and will in general describe all the important charge polarization effects.

If methods including electron correlation are used, higher angular momentum functions are essential. Electron correlation describes the energy lowering by the electrons "avoiding" each other, beyond the average effect taken into account by HF methods. Two types of correlation can be identified, an "in-out" and an "angular" correlation. The "in-out" or *radial correlation* refers to the situation where one electron is close to, and the other far from, the nucleus. To describe this, the basis set needs functions of the same type, but with different exponents. The *angular correlation* refers to the situation where two electrons are on opposite sides of the nucleus. To describe this, the basis set needs functions of same magnitude exponents, but different angular momenta. For example, to describe the angular correlation of an s-function, p-functions (and d-, f-, g-functions etc.) are needed. The angular correlation is of similar importance as the radial correlation, and higher angular momentum functions are consequently essential for correlated calculations. Although these properly should be labelled correlation functions, they also serve as polarization functions for HF wave functions, and it is common to denote them as polarization functions. Normally only the correlation of the valence electrons is considered, and the exponent of the polarization functions should be of the same magnitude as the valence s- and p-functions. In contrast to HF methods, the higher angular momentum functions (beyond the first set of polarization functions) are quite important. Or alternatively formulated, the convergence in terms of angular momentum is slower for correlated wave functions than at the HF level. For a basis set which is complete up to angular momentum L , numerical analysis suggests that the asymptotic convergence at the HF level is exponential (i.e. $\sim \exp(-L)$), while it is $\sim (L+1)^{-3}$ at correlated levels.⁴

Polarization functions are added to the chosen sp-basis. Adding a single set of polarization functions (p-functions on hydrogens and d-functions on heavy atoms) to the DZ basis forms a *Double Zeta plus Polarization* (DZP) type basis. There is a variation where polarization functions are only added to non-hydrogen atoms. This does not mean that polarization functions are not important on hydrogens. However, hydrogens often have a "passive" role, sitting at the end of bond which does not take an active part in the property of interest. The errors introduced by not including hydrogen polarization functions are often rather constant and, as the interest is usually in energy differences,

they tend to cancel out. As hydrogens often account for a large number of atoms in the system, a saving of three basis functions for each hydrogen is significant. If hydrogens play an important role in the property of interest, it is of course not a good idea to neglect polarization functions on hydrogens.

Similarly to the sp-basis sets, multiple sets of polarization functions with different exponents may be added. If two sets of polarization functions are added to a TZ sp-basis, a *Triple Zeta plus Double Polarization* (TZ2P) type basis is obtained. For larger basis sets with many polarization functions the explicit composition in terms of number and types of functions is usually given. At the HF level there is usually little gained by expanding the basis set beyond TZ2P, and even a DZP type basis set usually gives "good" results (compared to the HF limit). Correlated methods, however, require more, and higher, angular momentum, polarization functions to achieve the same level of convergence.

Before moving on we need to introduce the concept of *basis set balance*. In principle many sets of polarization functions may be added to a small sp-basis. This is not a good idea. If an insufficient number of sp-functions has been chosen for describing the fundamental electron distribution, the optimization procedure used in obtaining the wave function (and possibly also the geometry) may try to compensate for inadequacies in the sp-basis by using higher angular momentum functions, producing artefacts. A rule of thumb says that the number of functions of a given type should at most be one less than the type with one lower angular momentum. A 3s2p1d basis is balanced, but a 3s2p2d2f1g basis is too heavily polarized. It may not be necessary to polarize the basis all the way up, thus a 5s4p3d2f1g basis is balanced, but if it is known (for example by comparison with experimental data) that f- and g-functions are unimportant, they may be left out. Furthermore, it may be that 2 d-functions are sufficient for the given purpose, although a 5s4p1d basis would be considered underpolarized.

Another aspect of basis set balance is the occasional use of *mixed* basis sets, for example a DZP quality on the atoms in the "interesting" part of the molecule and a minimum basis for the "spectator" atoms. Another example would be addition of polarization functions for only a few hydrogens which are located "near" the reactive part of the system. For a large molecule this may lead to a substantial saving in the number of basis functions. It should be noted that this may bias the results and can create artefacts. For example, a calculation on the H₂ molecule with a minimum basis at one end and a DZ basis at the other end will predict that H₂ has a dipole moment, since the variational principle will preferentially place the electrons near the centre with the most basis functions. The majority of calculations are therefore performed with basis sets of the same quality (minimum, DZP, TZ2P, ...) on all atoms, possibly cutting polarization and/or diffuse (small exponent) functions on hydrogens. Even so, it may be argued that small basis sets inherently tend to be unbalanced. Consider for example the LiF molecule in a minimum or DZ type basis. This will have a very ionic structure, Li⁺F⁻, with nearly all the valence electrons being located at the fluorine. In terms of number of basis functions per electron, the Li basis is thus of a much higher quality than the F basis, and thereby unbalanced. Of course this effect diminishes as the size of the atomic basis set increases.

Except for very small systems it is impractical to saturate the basis set so that the absolute error in the energy is reduced below chemical accuracy, for example 1 kcal/mol. The important point in choosing a balanced basis set is to keep the error as constant

as possible. The use of mixed basis sets should therefore only be done after careful consideration. Furthermore, the use of small basis sets for systems containing elements with substantially different numbers of valence electrons (like LiF) may produce artefacts.

Having decided on the number of basis functions (from a consideration of the property of interest and the computational cost), the question becomes: how are the values for the exponents in the basis functions chosen? The values for s- and p-functions are typically determined by performing variational HF calculations for atoms, using the exponents as variational parameters. The exponent values which give the lowest energy are the "best", at least for the atom. In some cases the optimum exponents are chosen on the basis of minimizing the energy of a wave function which includes electron correlation. The HF procedure cannot be used for determining exponents of polarization functions for atoms. By definition these functions are unoccupied in atoms, and therefore make no contribution to the energy. Suitable polarization exponents may be chosen by performing variational calculations on molecular systems (where the HF energy does depend on polarization functions) or on atoms with correlated wave functions. Since the main function of higher angular momentum functions is to recover electron correlation, the latter approach is usually preferred. Often only the optimum exponent is determined for a single polarization function, and multiple polarization functions are generated by splitting the exponents symmetrically around the optimum value for a single function. The splitting factor is typically taken in the range 2–4. For example if a single d-function for carbon has an exponent value of 0.8, two polarization functions may be assigned with exponents of 0.4 and 1.6 (splitting factor of 4).

5.3 Even- and Well-tempered Basis Sets

The optimization of basis function exponents is an example of a highly non-linear optimization (Chapter 14). When the basis set becomes large, the optimization problem is no longer easy. The basis functions start to become linearly dependent (the basis approaches completeness) and the energy becomes a very flat function of the exponents. Furthermore, the multiple local minima problem is encountered. An analysis of basis sets which have been optimized by variational methods reveals that the ratio between two successive exponents is approximately constant. Taking this ratio to be constant reduces the optimization problem to only two parameters for each type of basis function, independently of the size of the basis. Such basis sets were labelled *even-tempered* basis sets, with the *i*th exponent given as $\zeta_i = \alpha\beta^i$ where α and β are fixed constants for a given type of function and nuclear charge. It was later discovered that the optimum α and β constants to a good approximation can be written as functions of the size of the basis set, M .⁵

$$\begin{aligned}\zeta_i &= \alpha\beta^i, \quad i = 1, 2, \dots, M \\ \ln(\ln \beta) &= b \ln M + b' \\ \ln \alpha &= a \ln(\beta - 1) + a'\end{aligned}\tag{5.3}$$

The constants a, a', b and b' depend only on the atom type and the type of function. Even-tempered basis sets have the advantage that it is easy to generate a sequence of basis sets which are guaranteed to converge towards a complete basis. This is useful if

the attempt is to extrapolate a given property to the basis set limit. The disadvantage is that the convergence is somewhat slow, and an explicitly optimized basis set of a given size will usually give a better answer than an even-tempered basis of the same size.

Even-tempered basis sets have the same ratio between exponents over the whole range. From chemical considerations it is usually preferable to cover the valence region better than the core region. This may be achieved by *well-tempered* basis sets.⁶ The idea is similar to the even-tempered basis sets, the exponents are generated by a suitable formula containing only a few parameters to be optimized. The exponents in a well-tempered basis of size M are generated as:

$$\zeta_i = \alpha\beta^{i-1} \left(1 + \gamma \left(\frac{i}{M} \right)^\delta \right), \quad i = 1, 2, \dots, M \quad (5.4)$$

The parameters α, β, γ and δ are optimized for each atom. The exponents are the same for all types of angular momentum functions, and s-, p- and d-functions (and higher angular momentum) consequently have the same radial part.

Optimization of basis sets is not something the average user need to worry about. Optimized basis sets of many different sizes and qualities are available either in the forms of tables, or built into the computer programs. The user "merely" has to select a suitable basis set. However, if the interest is in specialized properties the basis set may need to be tailored to meet the specific needs. For example if the property of interest is an accurate value for the electron density at the nucleus (for example for determining the Fermi contact contribution to spin-spin coupling, see section 10.7.2) then basis functions with very large exponents are required. Alternatively, for calculating hyperpolarizabilities very diffuse functions are required. In such cases the basis function optimization is in terms of the property of interest, and not in terms of energy. Basis functions are added until the change upon addition of one extra function is less than a given threshold.

5.4 Contracted Basis Sets

One disadvantage of all energy optimized basis sets is the fact that they primarily depend on the wave function in the region of the inner shell electrons. The 1s-electrons account for a large part of the total energy, and minimizing the energy will tend to make the basis set optimum for the core electrons, and less than optimum for the valence electrons. However, chemistry is mainly dependent on the valence electrons. Furthermore, many properties (for example polarizability) depend mainly on the wave function "tail" (far from the nucleus), which energetically is unimportant. An energy optimized basis set which gives a good description of the outer part of the wave function therefore needs to be very large, with the majority of the functions being used to describe the 1s-electrons with an accuracy comparable to that for the outer electrons in an energetic sense. This is not the most efficient way of designing basis sets for describing the outer part of the wave function. Instead energy optimized basis sets are usually augmented explicitly with *diffuse* functions (basis functions with small exponents). Diffuse functions are needed whenever loosely bound electrons are present (for example in anions or excited states) or when the property of interest is dependent on the wave function tail (for example polarizability).

The fact that many basis functions go into describing the energetically important, but chemically unimportant, core electrons is the foundation for *contracted* basis sets. Consider for example a basis set consisting of 10 s-functions (and some p-functions) for carbon. Having optimized these 10 exponents by variational calculations on the carbon atom, maybe six of the 10 functions are found primarily to be used for describing the 1s-orbital, and two of the four remaining describe the “inner” part of the 2s-orbital. The important chemical region is the outer valence. Out of the 10 functions, only two are actually used for describing the chemically interesting phenomena. Considering that the computational cost increases as the fourth power (or higher) of the number of basis functions, this is very inefficient. As the core orbitals change very little depending on the chemical bonding situation, the MO expansion coefficients in front of these inner basis functions also change very little. The majority of the computational effort is therefore spent describing the chemically uninteresting part of the wave function, which furthermore is almost constant.

Consider now making the variational coefficients in front of the inner basis functions constant, i.e. they are no longer parameters to be determined by the variational principle. The 1s-orbital is thus described by a fixed linear combination of say six basis functions. Similarly the remaining four basis functions may be contracted into only two functions, for example by fixing the coefficient in front of the inner three functions. In doing this the number of basis functions to be handled by the variational procedure has been reduced from 10 to three.

Combining the full set of basis functions, known as the *primitive* GTOs (PGTOs), into a smaller set of functions by forming fixed linear combinations is known as basis set *contraction*, and the resulting functions are called *contracted* GTOs (CGTOs).

$$\chi(\text{CGTO}) = \sum_i^k a_i \chi_i(\text{PGTO}) \quad (5.3)$$

The previously introduced acronyms DZP, TZ2P etc., refer to the number of contracted basis functions. Contraction is especially useful for orbitals describing the inner (core) electrons, since they require a relatively large number of functions for representing the wave function cusp near the nucleus, and furthermore are largely independent of the environment. Contracting a basis set will always increase the energy, since it is a restriction of the number of variational parameters, and makes the basis set less flexible, but will also reduce the computational cost significantly. The decision is thus how much loss in accuracy is acceptable compared to the gain in computational efficiency.

The *degree of contraction* is the number of PGTOs entering the CGTO, typically varying between 1 and 10. The specification of a basis set in terms of primitive and contracted functions is given by the notation (10s4p1d/4s1p) \rightarrow [3s2p1d/2s1p]. The basis in parentheses is the number of primitives with heavy atoms (first row elements) before the slash and hydrogen after. The basis in the square brackets is the number of contracted functions. Note that this does not tell how the contraction is done, it only indicates the size of the final basis (and thereby the size of the variational problem in HF calculations).

There are two different ways of contracting a set of primitive GTOs to a set of contracted GTOs: *segmented* and *general* contraction. Segmented contraction is the

oldest method, and the one used in the above example. A given set of PGTOs is partitioned into smaller sets of functions which are made into CGTO by determining suitable coefficients. A 10s basis may be contracted to 3s by taking the inner six functions as one CGTO, the next three as the second CGTO and the one remaining PGTO as the third "contracted" GTO.

$$\begin{aligned}\chi_1(\text{CGTO}) &= \sum_{i=1}^6 a_i \chi_i(\text{PGTO}) \\ \chi_2(\text{CGTO}) &= \sum_{i=7}^9 a_i \chi_i(\text{PGTO}) \\ \chi_3(\text{CGTO}) &= \chi_{10}(\text{PGTO})\end{aligned}\tag{5.4}$$

In a segmented contraction each primitive as a rule is only used in one contracted function. In some cases it may be necessary to duplicate one or two PGTOs in two adjacent CGTOs. The contraction coefficients can be determined by a variational optimization, for example from an atomic HF calculation.

In a general contraction all primitives (on a given atom) and of a given angular momentum enter all the contracted functions having that angular momentum, but with different contraction coefficients.

$$\begin{aligned}\chi_1(\text{CGTO}) &= \sum_{i=1}^{10} a_i \chi_i(\text{PGTO}) \\ \chi_2(\text{CGTO}) &= \sum_{i=1}^{10} b_i \chi_i(\text{PGTO}) \\ \chi_3(\text{CGTO}) &= \sum_{i=1}^{10} c_i \chi_i(\text{PGTO})\end{aligned}\tag{5.5}$$

One popular way of obtaining general contraction coefficients is from Atomic Natural Orbitals (ANO), to be discussed in section 5.4.4. The difference between segmented and general contraction may be illustrated as shown in Figure 5.2.

There are many different contracted basis sets available in the literature or built into programs, and the average user usually only needs to select a suitable quality basis for the calculation. Below is a short description of some basis sets which often are used in routine calculations.

5.4.1 Pople Style Basis Sets

STO-nG basis sets Slater Type Orbital consisting of n PGTOs.⁷ This is a minimum type basis where the exponents of the PGTO are determined by fitting to the STO, rather than optimizing them by a variational procedure. Although basis sets with $n = 2-6$ have been derived, it has been found that using more than three PGTOs to represent the STO gives little improvement, and the STO-3G basis is a widely used minimum basis. This type of basis set has been determined for many elements of the periodic table. The designation of the carbon/hydrogen STO-3G basis is (6s3p/3s) -- [2s1p/1s].

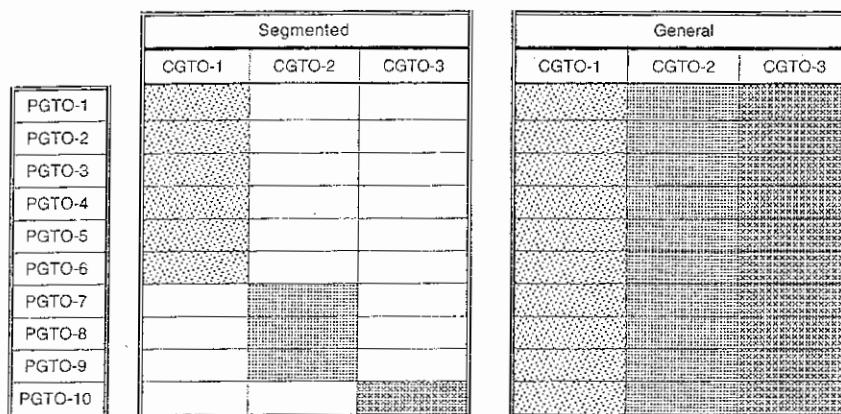


Figure 5.2 Segmented and general contraction

k-nlmG basis sets These basis sets have been designed by Pople and co-workers, and are of the split valence type, with the k in front of the dash indicating how many PGTOs are used for representing the core orbitals. The nlm after the dash indicate both how many functions the valence orbitals are split into, and how many PGTOs are used for their representation. Two values (e.g. nl) indicate a split valence, while three values (e.g. nlm) indicate a triple split valence. The values before the G (for Gaussian) indicate the s - and p -functions in the basis; the polarization functions are placed after the G. This type of basis sets has the further restriction that the same exponent is used for both the s - and p -functions in the valence. This increases the computational efficiency, but of course decreases the flexibility of the basis set. The exponents in the PGTO have been optimized by variational procedures.

3-21G This is a split valence basis, where the core orbitals are a contraction of three PGTOs, the inner part of the valence orbitals is a contraction of two PGTOs and the outer part of the valence is represented by one PGTO.⁸ The designation of the carbon/hydrogen 3-21G basis is $(6s3p/3s) \rightarrow [3s2p/2s]$. Note that the 3-21G basis contains the same number of primitive GTOs as the STO-3G, however, it is much more flexible as there are twice as many valence functions which can combine freely to make MOs.

6-31G This is also a split valence basis, where the core orbitals are a contraction of six PGTOs, the inner part of the valence orbitals is a contraction of three PGTOs and the outer part of the valence represented by one PGTO.⁹ The designation of the carbon/hydrogen 6-31G basis is $(10s4p/4s) \rightarrow [3s2p/2s]$. In terms of contracted basis functions it contains the same number as 3-21G, but the representation of each functions is better since more PGTOs are used.

6-311G This is a triple split valence basis, where the core orbitals are a contraction of six PGTOs and the valence split into three functions, represented by three, one, and one PGTOs, respectively.¹⁰

To each of these basis sets can be added diffuse¹¹ and/or polarization functions. Diffuse functions are normally s- and p-functions and consequently go before the G. They are denoted by + or ++, with the first + indicating one set of diffuse s- and p-functions on heavy atoms, and the second + indicating that a diffuse s-function is also added to hydrogens. The arguments for adding only diffuse functions on non-hydrogen atoms is the same as that for adding only polarization functions on non-hydrogens (Section 5.2). Polarization functions are indicated after the G, with a separate designation for heavy atoms and hydrogens. The 6-31+G(d) is a split valence basis with one set of diffuse sp-functions on heavy atoms only and a single d-type polarization function on heavy atoms. A 6-311++G(2df,2pd) is similarly a triple split valence with additional diffuse sp-functions, and two d- and one f-functions on heavy atoms and diffuse s- and two p- and one d-functions on hydrogens. The largest standard Pople style basis set is 6-311++G(3df, 3pd). These types of basis sets have been derived for hydrogen and the first row elements, and some of the basis sets have also been derived for second and higher row elements.

If only one set of polarization functions is used, an alternative notation in terms of * is also widely used. The 6-31G* basis is identical to 6-31G(d), and 6-31G** is identical to 6-31G(d,p). A special note should be made for the 3-21G* basis. The 3-21G basis is basically too small to support polarization functions (it becomes unbalanced). However, the 3-21G basis by itself performs poorly for hypervalent molecules, such as sulfoxides and sulfones. This can be substantially improved by adding a set of d-functions. The 3-21G* basis has only d-functions on second row elements (it is sometimes denoted 3-21G(*) to indicate this), and should not be considered a polarized basis. Rather, the addition of a set of d-functions should be considered an *ad hoc* repair of a known flaw.

5.4.2 Dunning–Huzinaga Basis Sets

Huzinaga determined uncontracted energy optimized basis sets up to (10s6p) for first row elements.¹³ This was later extended to (14s9p) by van Duijneveldt¹⁴ and up to (18s, 13p) by Partridge.¹⁵ Dunning used the Huzinaga primitive GTOs to derive various contraction schemes (DH type basis sets).¹⁶ A DZ type basis can be made by a contraction of the (9s5p/4s) PGTO to [4s2p/2s]. The contraction scheme is 6,1,1,1 for s-functions, 4,1 for the p-functions, and 3,1 for hydrogen. A widely used split valence type basis is a contraction of the same primitive set to [3s2p/2s] where the s-contraction is 7,2,1 (note that one primitive enters twice). A widely used TZ type basis (actually only a triple split valence) is a contraction of the (10s6p/5s) to [5s3p/3s], with the contraction scheme 5,3,1,1,1 for s-functions, 4,1,1 for p-functions, and 3,1,1 for hydrogens. Again a duplication of one of the s-primitives has been allowed.

McLean and Chandler developed a similar set of contracted basis sets from Huzinaga's primitive optimized set for second row elements.¹⁷ A DZ type basis is derived by contracting (12s8p) → [5s3p], and a TZ type is derived by contracting (13s9p) → [6s4p]. The latter contraction is 6,3,1,1,1 for the s-functions and 4,2,1,1,1 for the p-functions, and is often used in connection with the Pople 6-31G when second row elements are present.

The Dunning–Huzinaga type basis sets do not have the restriction of the Pople style basis sets of equal exponents for the s- and p-functions, and they are therefore somewhat more flexible, but computationally also more expensive. The major determining factor,

however, is the number of basis functions and less the exact description of each function. Normally there is little difference in the performance of different DZ or different TZ type basis sets.

The primary reason for the popularity of the Pople and DH style basis sets is the extensive calibration available. There have been so many calculations reported with these basis sets that it is possible to get a fairly good idea of the level of accuracy that can be attained with a given basis. This is of course a self-sustaining procedure, the more calculations that are reported with a given basis, the more popular it becomes, since the calibration set becomes larger and larger.

5.4.3 MINI, MIDI and MAXI basis sets

Tatewaki and Huzinaga optimized minimum basis sets for a large part of the periodic system.¹⁸ The MINI- i ($i = 1-4$) basis are all minimum basis sets with three PGTOs in the 2s CGTO and varying numbers of PGTOs in the 1s and 2p CGTOs. In terms of PGTOs the MINI-1 is (3s,3s,3p), MINI-2 is (3s,3s,4p), MINI-3 is (4s,3s,3p) and MINI-4 is (4s,3s,4p). These MINI basis sets in general perform better than STO-3G, but it should be kept in mind that they still are minimum basis sets. The MIDI- i basis sets are identical to MINI- i , except that the outer valence function is decontracted. The MAXI- i basis sets all employ four PGTOs for the 2s CGTO and between five and seven PGTOs for the 1s and 2p CGTOs. The valence orbitals are split into three or four functions, and MAXI-1 is (9s5p) \rightarrow [4s3p] (contraction 5,2,1,1 and 3,1,1), MAXI-3 (10s6p) \rightarrow [5s4p] (contraction 6,2,1,1,1 and 3,1,1,1) and MAXI-5 is (11s7p) \rightarrow [5s4p] (contraction 7,2,1,1,1 and 4,1,1,1).

5.4.4 Atomic Natural Orbitals Basis Sets

All of the above basis sets are of the segmented contraction type. Modern contracted basis sets aimed at producing very accurate wave functions often employ a general contraction scheme. The ANO and cc basis sets below are of the general contraction type.

The idea in the *Atomic Natural Orbitals* (ANO) type basis sets is to contract a large PGTO set to a fairly small number of CGTOs by using *natural orbitals* from a correlated calculation on the free atom, typically at the CISD level.¹⁹ The natural orbitals are those which diagonalize the density matrix, and the eigenvalues are called *orbital occupation numbers* (see Section 9.5). The orbital occupation number is the number of electrons in the orbital. For an RHF wave function, ANOs would be identical to the canonical orbitals with occupation numbers of exactly 0 or 2. When a correlated wave function is used, however, the occupation number may have any value between 0 and 2. The ANO contraction selects the important combinations of the PGTOs from the magnitude of the occupation numbers. A large primitive basis, typically generated as an even-tempered sequence, may generate several different contracted basis sets by gradually lowering the selection threshold for the occupation number. The nice feature of the ANO contraction is that it more or less “automatically” generates balanced basis sets, e.g. for neon the ANO procedure generates the following basis set: [2s1p], [3s2p1d], [4s3p2d1f] and [5s4p3d2f1g]. Furthermore, in such a sequence the smaller ANO basis sets are true subsets of the larger, since the same primitive set of functions is used.

The Theoretical Background

Ab initio molecular orbital theory is concerned with predicting the properties of atomic and molecular systems. It is based upon the fundamental laws of quantum mechanics and uses a variety of mathematical transformation and approximation techniques to solve the fundamental equations. This appendix provides an introductory overview of the theory underlying ab initio electronic structure methods. The final section provides a similar overview of the theory underlying Density Functional Theory methods.

The Schrödinger Equation

Quantum mechanics explains how entities like electrons have both particle-like and wave-like characteristics. The Schrödinger equation describes the wavefunction of a particle:

$$\left\{ \frac{-\hbar^2}{8\pi^2 m} \nabla^2 + V \right\} \Psi(\mathbf{r}, t) = \frac{i\hbar}{2\pi} \frac{\partial \Psi(\mathbf{r}, t)}{\partial t} \quad [1]$$

In this equation, Ψ is the wavefunction, m is the mass of the particle, \hbar is Planck's constant, and V is the potential field in which the particle is moving.[†] The product of Ψ with its complex conjugate ($\Psi^*\Psi$, often written as $|\Psi|^2$) is interpreted as the probability distribution of the particle.

The Schrödinger equation for a collection of particles like a molecule is very similar. In this case, Ψ would be a function of the coordinates of all the particles in the system as well as t .

The energy and many other properties of the particle can be obtained by solving the Schrödinger equation for Ψ , subject to the appropriate boundary conditions. Many different wavefunctions are solutions to it, corresponding to different stationary states of the system.

[†] The differential operator on the left side of the equation is known as "del-squared." The operator del is equivalent to partial differentiation with respect to x , y , and z components:

$$\nabla = \frac{\partial}{\partial x} \mathbf{i} + \frac{\partial}{\partial y} \mathbf{j} + \frac{\partial}{\partial z} \mathbf{k}$$

If V is not a function of time, the Schrödinger equation can be simplified using the mathematical technique known as separation of variables. If we write the wavefunction as the product of a spatial function and a time function:

$$\Psi(\hat{\mathbf{r}}, t) = \psi(\hat{\mathbf{r}})\tau(t) \quad [2]$$

and then substitute these new functions into Equation 1, we will obtain two equations, one of which depends on the position of the particle independent of time and the other of which is a function of time alone. For the problems in which we are interested, this separation is valid, and we focus entirely on the familiar time-independent Schrödinger equation:

$$\mathbf{H}\psi(\hat{\mathbf{r}}) = E\psi(\hat{\mathbf{r}}) \quad [3]$$

where E is the energy of the particle, and \mathbf{H} is the *Hamiltonian operator*, equal to:

$$\mathbf{H} = \frac{-\hbar^2}{8\pi^2 m} \nabla^2 + V \quad [4]$$

The various solutions to Equation 3 correspond to different stationary states of the particle (molecule). The one with the lowest energy is called the *ground state*. Equation 3 is a non-relativistic description of the system which is not valid when the velocities of particles approach the speed of light. Thus, Equation 3 does not give an accurate description of the core electrons in large nuclei.

Note also that Equation 3 is an *eigenvalue equation*: an equation in which an operator acting on a function produces a multiple of the function itself as its result, having the general form:

$$\mathbf{Op} f = c f \quad [5]$$

where \mathbf{Op} is an operator, f is a function, and c is a constant. The set of functions for which the equation holds are its *eigenfunctions*, each of which has an associated value for c , known as its *eigenvalue*. In the case of the Schrödinger equation, the eigenvalues are the energies corresponding to the different stationary states of the molecular system.

The Molecular Hamiltonian

For a molecular system, Ψ is a function of the positions of the electrons and the nuclei within the molecule, which we will designate as \mathbf{r} and \mathbf{R} , respectively. These symbols are a shorthand for the set of component vectors describing the position of each particle. We'll use subscripted versions of them to denote the vector corresponding to a particular electron or nucleus: \mathbf{r}_i and \mathbf{R}_J . Note that electrons are treated individually, while each nucleus is treated as an aggregate; the component nucleons are not treated individually.

The Hamiltonian is made up of kinetic and potential energy terms:

$$\mathbf{H} = \mathbf{T} + \mathbf{V} \quad [6]$$

The kinetic energy is a summation of ∇^2 over all the particles in the molecule:

$$\mathbf{T} = -\frac{\hbar^2}{8\pi^2} \sum_k \frac{1}{m_k} \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) \quad [7]$$

The potential energy component is the Coulomb repulsion between each pair of charged entities (treating each atomic nucleus as a single charged mass):

$$\mathbf{V} = \frac{1}{4\pi\epsilon_0} \sum_j \sum_{k < j} \frac{e_j e_k}{\Delta r_{jk}} \quad [8]$$

where Δr_{jk} is the distance between the two particles, and e_j and e_k are the charges on particles j and k . For an electron, the charge is $-e$, while for a nucleus, the charge is Ze , where Z is the atomic number for that atom. Thus,

$$\mathbf{V} = \frac{1}{4\pi\epsilon_0} \left(- \sum_i \sum_l \left(\frac{Z_l e^2}{\Delta r_{il}} \right) + \sum_i \sum_{j < i} \left(\frac{e^2}{\Delta r_{ij}} \right) + \sum_l \sum_{J < l} \left(\frac{Z_l Z_J e^2}{\Delta R_{lJ}} \right) \right) \quad [9]$$

The first term corresponds to electron-nuclear attraction, the second to electron-electron repulsion, and the third to nuclear-nuclear repulsion.



Atomic Units

The fundamental equations of quantum chemistry are usually expressed in units designed to simplify their form by eliminating fundamental constants. The atomic unit of length is the *Bohr radius*:

$$a_0 = \frac{h^2}{4\pi^2 m_e e^2} = 0.52917725 \text{ \AA} \quad [10]$$

Coordinates can be transformed to bohrs by dividing them by a_0 . Energies are measured in *hartrees*, defined as the Coulomb repulsion between two electrons separated by 1 bohr:

$$1 \text{ hartree} = \frac{e^2}{a_0} \quad [11]$$

Masses are also specified in terms of electron mass units (i.e. define $m_e=1$).

We will use these units in all future equations.

The Born-Oppenheimer Approximation

The Born-Oppenheimer approximation is the first of several approximations used to simplify the solution of the Schrödinger equation. It simplifies the general molecular problem by separating nuclear and electronic motions. This approximation is reasonable since the mass of a typical nucleus is thousands of times greater than that of an electron. The nuclei move very slowly with respect to the electrons, and the electrons react essentially instantaneously to changes in nuclear position. Thus, the electron distribution within a molecular system depends on the positions of the nuclei, and not on their velocities. Put another way, the nuclei look fixed to the electrons, and electronic motion can be described as occurring in a field of fixed nuclei.

The full Hamiltonian for the molecular system can then be written as:

$$\mathbf{H} = \mathbf{T}^{elec}(\hat{\mathbf{r}}) + \mathbf{T}^{nucl}(\hat{\mathbf{R}}) + \mathbf{V}^{nucl-elec}(\hat{\mathbf{R}}, \hat{\mathbf{r}}) + \mathbf{V}^{elec}(\hat{\mathbf{r}}) + \mathbf{V}^{nucl}(\hat{\mathbf{R}}) \quad [12]$$

The Born-Oppenheimer approximation allows the two parts of the problem to be solved independently, so we can construct an electronic Hamiltonian which neglects the kinetic energy term for the nuclei:

$$\begin{aligned}
 \mathbf{H}^{elec} = & -\frac{1}{2} \sum_i^{electrons} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) - \sum_i \sum_I^{electrons\ nuclei} \left(\frac{Z_I}{|\vec{\mathbf{R}}_I - \vec{\mathbf{r}}_i|} \right) \\
 & + \sum_i \sum_{j < i}^{electrons} \left(\frac{1}{|\vec{\mathbf{r}}_i - \vec{\mathbf{r}}_j|} \right) + \sum_I \sum_{J < I}^{nuclei} \left(\frac{Z_I Z_J}{|\vec{\mathbf{R}}_I - \vec{\mathbf{R}}_J|} \right)
 \end{aligned}
 \tag{13}$$

Note that the fundamental physical constants drop out with the use of atomic units.

This Hamiltonian is then used in the Schrödinger equation describing the motion of electrons in the field of fixed nuclei:

$$\mathbf{H}^{elec} \psi^{elec}(\vec{\mathbf{r}}, \vec{\mathbf{R}}) = E^{eff}(\vec{\mathbf{R}}) \psi^{elec}(\vec{\mathbf{r}}, \vec{\mathbf{R}})
 \tag{14}$$

Solving this equation for the electronic wavefunction will produce the effective nuclear potential function E^{eff} .[†] It depends on the nuclear coordinates and describes the potential energy surface for the system.

Accordingly, E^{eff} is also used as the effective potential for the nuclear Hamiltonian:

$$\mathbf{H}^{nucl} = \mathbf{T}^{nucl}(\vec{\mathbf{R}}) + E^{eff}(\vec{\mathbf{R}})
 \tag{15}$$

This Hamiltonian is used in the Schrödinger equation for nuclear motion, describing the vibrational, rotational, and translational states of the nuclei. Solving the nuclear Schrödinger equation (at least approximately) is necessary for predicting the vibrational spectra of molecules.

From this point on, we will focus entirely on the electronic problem. We will omit the superscripts on all operators and functions.

Restrictions on the Wavefunction

We've noted that ψ^2 is interpreted as the probability density for the particle(s) it describes. Therefore, we require that ψ be *normalized*; if we integrate over all space,

[†] For a given set of nuclear coordinates, this corresponds to the total energy predicted by a single point energy calculation, although such calculations, of course, do not solve this equation exactly. The approximation methods used to solve it will be discussed in subsequent sections of this appendix.

the probability should be the number of particles (the particles *are* somewhere). Accordingly, we multiply ψ by a constant such that:

$$\int_{-\infty}^{+\infty} |c\psi|^2 dv = n_{\text{particles}} \quad \dagger \quad [16]$$

We can do this because the Schrödinger equation is an eigenvalue equation, and in general, if f is a solution to an eigenvalue equation, then cf is also, for any value of c . For the Schrödinger equation, it is easy to show that $H(c\psi) = cH(\psi)$ and that $E(c\psi) = c(E\psi)$; thus, if ψ is a solution to the Schrödinger equation, then $c\psi$ is as well.

Secondly, ψ must also be *antisymmetric*, meaning that it must change sign when two identical particles are interchanged. For a simple function, antisymmetry means that the following relation holds:

$$f(i, j) = -f(j, i) \quad [17]$$

For an electronic wavefunction, antisymmetry is a physical requirement following from the fact that electrons are fermions.[‡] It is essentially a requirement that ψ agree with the results of experimental physics. More specifically, this requirement means that any valid wavefunction must satisfy the following condition:

$$\psi(\hat{\mathbf{r}}_1, \dots, \hat{\mathbf{r}}_i, \dots, \hat{\mathbf{r}}_j, \dots, \hat{\mathbf{r}}_i, \dots, \hat{\mathbf{r}}_j) = -\psi(\hat{\mathbf{r}}_1, \dots, \hat{\mathbf{r}}_j, \dots, \hat{\mathbf{r}}_i, \dots, \hat{\mathbf{r}}_i) \quad [18]$$

Hartree-Fock Theory

The underlying physical laws necessary for the mathematical theory of a large part of physics and the whole of chemistry are thus completely known, and the difficulty is only that the exact application of these laws leads to equations much too complicated to be soluble.

— P. A. M. Dirac, 1929

[†] If ψ is complex, the integral becomes: $c^*c \iiint_{-\infty}^{+\infty} \psi^* \psi dx dy dz = n$

[‡] Fermions are particles that have the properties of antisymmetry and a half-integral spin quantum number, among others.

An exact solution to the Schrödinger equation is not possible for any but the most trivial molecular systems. However, a number of simplifying assumptions and procedures do make an approximate solution possible for a large range of molecules.

Molecular Orbitals

The first approximation we'll consider comes from the interpretation of $|\psi|^2$ as a probability density for the electrons within the system. Molecular orbital theory decomposes ψ into a combination of molecular orbitals: ϕ_1, ϕ_2, \dots . To fulfill some of the conditions on ψ we discussed previously, we choose a normalized, orthogonal set of molecular orbitals:

$$\begin{aligned} \iiint \phi_i^* \phi_i dx dy dz &= 1 \\ \iiint \phi_i^* \phi_j dx dy dz &= 0; \quad i \neq j \end{aligned} \quad [19]$$

The simplest possible way of making ψ as a combination of these molecular orbitals is by forming their *Hartree product*:

$$\psi(\vec{r}) = \phi_1(\vec{r}_1) \phi_2(\vec{r}_2) \dots \phi_n(\vec{r}_n) \quad [20]$$

However, such a function is not antisymmetric, since interchanging two of the \vec{r}_i 's—equivalent to swapping the orbitals of two electrons—does not result in a sign change. Hence, this Hartree product is an inadequate wavefunction.

Electron Spin

The simplest antisymmetric function that is a combination of molecular orbitals is a determinant. Before forming it, however, we need to account for a factor we've neglected so far: electron spin. Electrons can have spin up (+1/2) or down (-1/2). Equation 20 assumes that each molecular orbital holds only one electron. However, most calculations are closed shell calculations, using doubly occupied orbitals, holding two electrons of opposite spin. For the moment, we will limit our discussion to this case.

We define two spin functions, α and β , as follows:

$$\begin{aligned} \alpha(\uparrow) &= 1 & \alpha(\downarrow) &= 0 \\ \beta(\uparrow) &= 0 & \beta(\downarrow) &= 1 \end{aligned} \quad [21]$$

The α function is 1 for a spin up electron, and the β function is 1 when the electron is spin down. The notation $\alpha(i)$ and $\beta(i)$ will designate the values of α and β for electron i ; thus, $\alpha(1)$ is the value of α for electron 1.

Multiplying a molecular orbital function by α or β will include electron spin as part of the overall electronic wavefunction ψ . The product of the molecular orbital and a spin function is defined as a *spin orbital*, a function of both the electron's location and its spin.[†] Note that these spin orbitals are also orthonormal when the component molecular orbitals are.

We can now build a closed shell wavefunction by defining $n/2$ molecular orbitals for a system with n electrons, and then assigning electrons to these orbitals in pairs of opposite spin:

$$\psi(\vec{r}) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\vec{r}_1)\alpha(1) & \phi_1(\vec{r}_1)\beta(1) & \phi_2(\vec{r}_1)\alpha(1) & \phi_2(\vec{r}_1)\beta(1) & \dots & \phi_{\frac{n}{2}}(\vec{r}_1)\alpha(1) & \phi_{\frac{n}{2}}(\vec{r}_1)\beta(1) \\ \phi_1(\vec{r}_2)\alpha(2) & \phi_1(\vec{r}_2)\beta(2) & \phi_2(\vec{r}_2)\alpha(2) & \phi_2(\vec{r}_2)\beta(2) & \dots & \phi_{\frac{n}{2}}(\vec{r}_2)\alpha(2) & \phi_{\frac{n}{2}}(\vec{r}_2)\beta(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_1(\vec{r}_i)\alpha(i) & \phi_1(\vec{r}_i)\beta(i) & \phi_2(\vec{r}_i)\alpha(i) & \phi_2(\vec{r}_i)\beta(i) & \dots & \phi_{\frac{n}{2}}(\vec{r}_i)\alpha(i) & \phi_{\frac{n}{2}}(\vec{r}_i)\beta(i) \\ \phi_1(\vec{r}_j)\alpha(j) & \phi_1(\vec{r}_j)\beta(j) & \phi_2(\vec{r}_j)\alpha(j) & \phi_2(\vec{r}_j)\beta(j) & \dots & \phi_{\frac{n}{2}}(\vec{r}_j)\alpha(j) & \phi_{\frac{n}{2}}(\vec{r}_j)\beta(j) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_1(\vec{r}_n)\alpha(n) & \phi_1(\vec{r}_n)\beta(n) & \phi_2(\vec{r}_n)\alpha(n) & \phi_2(\vec{r}_n)\beta(n) & \dots & \phi_{\frac{n}{2}}(\vec{r}_n)\alpha(n) & \phi_{\frac{n}{2}}(\vec{r}_n)\beta(n) \end{vmatrix} \quad [22]$$

Each row is formed by representing all possible assignments of electron i to all orbital-spin combinations. The initial factor is necessary for normalization. Swapping two electrons corresponds to interchanging two rows of the determinant, which will have the effect of changing its sign.

This formulation is not just a mathematical trick to form an antisymmetric wavefunction. Quantum mechanics specifies that an electron's location is not deterministic but rather consists of a probability density; in this sense, it can be anywhere. This determinant mixes all of the possible orbitals of all of the electrons in the molecular system to form the wavefunction.

[†]Some texts use a separate notation for spin orbitals. We will not do so here.

Basis Sets

The next approximation involves expressing the molecular orbitals as linear combinations of a pre-defined set of one-electron functions known as basis functions. These basis functions are usually centered on the atomic nuclei and so bear some resemblance to atomic orbitals. However, the actual mathematical treatment is more general than this, and any set of appropriately defined functions may be used.

An individual molecular orbital is defined as:

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} \quad [23]$$

where the coefficients $c_{\mu i}$ are known as the *molecular orbital expansion coefficients*. The basis functions $\chi_1 \dots \chi_N$ are also chosen to be normalized. We follow the usual notational convention of using roman subscripts on molecular orbital functions and Greek subscripts on basis functions. Thus, χ_{μ} refers to an arbitrary basis function in the same way that ϕ_i refers to an arbitrary molecular orbital.

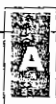
Gaussian and other ab initio electronic structure programs use gaussian-type atomic functions as basis functions. Gaussian functions have the general form:

$$g(\alpha, \hat{\mathbf{r}}) = c x^n y^m z^l e^{-\alpha r^2} \quad [24]$$

where $\hat{\mathbf{r}}$ is of course composed of x , y and z . α is a constant determining the size (radial extent) of the function. In a gaussian function, $e^{-\alpha r^2}$ is multiplied by powers (possibly 0) of x , y , and z , and a constant for normalization, so that:

$$\int_{\text{all space}} g^2 = 1 \quad [25]$$

Thus, c depends on α , l , m and n .



Here are three representative gaussian functions (s, p_y and d_{xy} types, respectively):

$$\begin{aligned} g_x(\alpha, \hat{\mathbf{r}}) &= \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} \\ g_y(\alpha, \hat{\mathbf{r}}) &= \left(\frac{128\alpha^5}{\pi^3}\right)^{1/4} y e^{-\alpha r^2} \\ g_{xy}(\alpha, \hat{\mathbf{r}}) &= \left(\frac{2048\alpha^7}{\pi^3}\right)^{1/4} xy e^{-\alpha r^2} \end{aligned} \quad [26]$$

Linear combinations of *primitive gaussians* like these are used to form the actual basis functions; the latter are called *contracted gaussians* and have the form:

$$\chi_\mu = \sum_p d_{\mu p} g_p \quad [27]$$

where the $d_{\mu p}$'s are fixed constants within a given basis set. Note that contracted functions are also normalized in common practice.

All of these constructions result in the following expansion for molecular orbitals:

$$\phi_i = \sum_\mu c_{\mu i} \chi_\mu = \sum_\mu c_{\mu i} \left(\sum_p d_{\mu p} g_p \right) \quad [28]$$

The Variational Principle

The problem has now become how to solve for the set of molecular orbital expansion coefficients, $c_{\mu i}$. Hartree-Fock theory takes advantage of the *variational principle*, which says that for the ground state of any antisymmetric normalized function of the electronic coordinates, which we will denote Ξ , then the expectation value for the energy corresponding to Ξ will always be greater than the energy for the *exact* wavefunction:

$$E(\Xi) > E(\Psi); \quad \Xi \neq \Psi \quad [29]$$

In other words, the energy of the exact wavefunction serves as a lower bound to the energies calculated by any other normalized antisymmetric function. Thus, the problem becomes one of finding the set of coefficients that minimize the energy of the resultant wavefunction.

The Roothaan-Hall Equations

The variational principle leads to the following equations describing the molecular orbital expansion coefficients, $c_{\nu i}$, derived by Roothaan and by Hall[†]:

$$\sum_{\nu=1}^N (F_{\mu\nu} - \epsilon_i S_{\mu\nu}) c_{\nu i} = 0 \quad \mu = 1, 2, \dots, N \quad [30]$$

Equation 30 can be rewritten in matrix form:

$$FC = SC\epsilon \quad [31]$$

where each element is a matrix. ϵ is a diagonal matrix of orbital energies, each of its elements ϵ_i is the one-electron orbital energy of molecular orbital χ_i .

F is called the *Fock matrix*, and it represents the average effects of the field of all the electrons on each orbital. For a closed shell system, its elements are:

$$F_{\mu\nu} = H_{\mu\nu}^{core} + \sum_{\lambda=1}^N \sum_{\sigma=1}^N P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\nu\sigma) \right] \quad [32]$$

where $H_{\mu\nu}^{core}$ is another matrix representing the energy of a single electron in the field of the bare nuclei, and P is the *density matrix*, defined as:

$$P_{\lambda\sigma} = 2 \sum_{i=1}^{occupied} c_{\lambda i}^* c_{\sigma i} \quad [33]$$

The coefficients are summed over the occupied orbitals only, and the factor of two comes from the fact that each orbital holds two electrons.

Finally, the matrix S from Equation 31 is the *overlap matrix*, indicating the overlap between orbitals.

Both the Fock matrix—through the density matrix—and the orbitals depend on the molecular orbital expansion coefficients. Thus, Equation 31 is not linear and must be solved iteratively. The procedure which does so is called the *Self-Consistent Field*

[†] We alter the subscripts slightly here from what has preceded in order to follow common usage.

(SCF) method. At convergence, the energy is at a minimum, and the orbitals generate a field which produces the same orbitals, accounting for the method's name. The solution produces a set of orbitals, both *occupied* (ϕ_i, \dots) and *virtual* (unoccupied, conventionally denoted $\phi_{a,b,\dots}$). The total number of orbitals is equal to the number of basis functions used.

The term $(\mu\nu|\lambda\sigma)$ in Equation 32 signifies the *two-electron repulsion integrals*. Under the Hartree-Fock treatment, each electron sees all of the other electrons as an average distribution; there is no instantaneous electron-electron interaction included. Higher level methods attempt to remedy this neglect of electron correlation in various ways, as we shall see.

The general strategy used by the SCF method (after initial setup steps) is as follows:

- ◆ Evaluate the integrals. In a conventional algorithm, they are stored on disk and read in for each iteration. In a direct algorithm, integrals are computed a few at a time as the Fock matrix is formed.
- ◆ Form an initial guess for the molecular orbital coefficients, and construct the density matrix.
- ◆ Form the Fock matrix.
- ◆ Solve for the density matrix.
- ◆ Test for convergence. If it fails, begin the next iteration. If it succeeds, go on to perform other parts of the calculation (such as population analysis).

Open Shell Methods

So far, we have considered only the restricted Hartree-Fock method. For open shell systems, an unrestricted method, capable of treating unpaired electrons, is needed.[†] For this case, the alpha and beta electrons are in different orbitals, resulting in two sets of molecular orbital expansion coefficients:

$$\begin{aligned}\phi_i^\alpha &= \sum_{\mu} c_{\mu i}^\alpha \chi_{\mu} \\ \phi_i^\beta &= \sum_{\mu} c_{\mu i}^\beta \chi_{\mu}\end{aligned}\quad [34]$$

[†] Refer also to the discussion of open shell calculations in Chapter 1 (page 10).

4 Electron Correlation Methods

The Hartree–Fock method generates solutions to the Schrödinger equation where the real electron–electron interaction is replaced by an average interaction (Chapter 3). In a sufficiently large basis, the HF wave function is able to account for $\sim 99\%$ of the total energy, but the remaining $\sim 1\%$ is often very important for describing chemical phenomena. The difference in energy between the HF and the lowest possible energy in a given basis set is called the *Electron Correlation* (EC) energy.¹ Physically it corresponds to the motion of the electrons being correlated, on average they are further apart than described by the HF wave function. As shown below, an UHF type of wave function is to a certain extent able to include electron correlation. The proper reference for discussing electron correlation is therefore a restricted (RHF or ROHF) wave function, although many authors use a UHF wave function for open-shell species. In the RHF case all the electrons are paired in molecular orbitals. The two electrons in a MO occupy the same physical space, and differ only in the spin function. The spatial overlap between the orbitals of two such “pair”-electrons is (exactly) one, while the overlap between two electrons belonging to different pairs is (exactly) zero, owing to the orthonormality of the MOs. This not the same as saying that there is no repulsion between electrons in different MOs, since the electron–electron repulsion integrals involve products of MOs ($\langle\phi_i|\phi_j\rangle = 0$ for $i \neq j$, but $\langle\phi_i\phi_j|\mathbf{g}|\phi_i\phi_j\rangle$ and $\langle\phi_i\phi_j|\mathbf{g}|\phi_j\phi_i\rangle$ are not necessarily zero).

Naively it may be expected that the correlation between pairs of electrons belonging to the same spatial MO would be the major part of the electron correlation. However, as the size of the molecule increases, the number of electron pairs belonging to different spatial MOs grows faster than those belonging to the same MO. Consider for example the valence orbitals for CH_4 . There are four intraorbital electron pairs of opposite spin, but there are 12 interorbital pairs of opposite spin, and 12 interorbital pairs of the same spin. A typical value for the intraorbital pair correlation of a single bond is ~ 20 kcal/mol, while that of an interorbital pair (where the two MO are spatially close, as in CH_4) is ~ 1 kcal/mol. The interpair correlation is therefore often comparable to the intrapair contribution.

Since the correlation between opposite spins has both intra- and inter-orbital contributions, it will be larger than the correlation between electrons having the same spin. The Pauli principle (or equivalently the antisymmetry of the wave function) has the consequence that there is no intraorbital correlation from electron pairs with the same spin. The opposite spin correlation is sometimes called the *Coulomb correlation*, while the same spin correlation is called the *Fermi correlation*, i.e. the Coulomb correlation is the largest contribution. Another way of looking at electron correlation is in terms of the electron density. In the immediate vicinity of an electron, there is a reduced probability of finding another electron. For electrons of opposite spin, this is often referred to as the *Coulomb hole*, the corresponding phenomenon for electrons of the same spin is the *Fermi hole*.

The HF method determines the best one-determinant trial wave function (within the given basis set). It is therefore clear that in order to improve on HF results, the starting point must be a trial wave function which contains more than one Slater Determinant (SD) Φ . This also means that the mental picture of electrons residing in orbitals has to be abandoned, and the more fundamental property, the electron density, should be considered. As the HF solution usually gives $\sim 99\%$ of the correct answer, electron correlation methods normally use the HF wave function as a starting point for improvements.

A generic multi-determinant trial wave function can be written as

$$\Psi = a_0\Phi_{\text{HF}} + \sum_{i=1} a_i\Phi_i \quad (4.1)$$

where a_0 usually is close to 1. Electron correlation methods differ in how they calculate the coefficients in front of the other determinants, a_0 being determined by the normalization condition.

As mentioned in Chapter 5, one can think of the expansion of an unknown MO in terms of basis functions as describing the MO “function” in the “coordinate system” of the basis functions. The multi-determinant wave function (4.1) can similarly be considered as describing the total wave function in a “coordinate” system of Slater determinants. The basis set determines the size of the one-electron basis (and thus limits the description of the one-electron functions, the MOs), while the number of determinants included determines the size of the many-electron basis (and thus limits the description of electron correlation).

4.1 Excited Slater Determinants

How are the additional determinants beyond the HF constructed? With N electrons and M basis functions, solution of the Roothaan–Hall equations for the RHF case will yield $N/2$ occupied MOs and $M - N/2$ unoccupied (virtual) MOs. Except for a minimum basis, there will always be more virtual than occupied MOs. A Slater determinant is determined by $N/2$ spatial MOs multiplied by two spin functions to yield N spinorbitals. By replacing MOs which are occupied in the HF determinant by MOs which are unoccupied, a whole series of determinants may be generated. These can be denoted according to how many occupied HF MOs have been replaced by unoccupied MOs, i.e. Slater determinants which are *singly*, *doubly*, *triply*, *quadruply* etc. *excited* relative to the HF determinant, up to a maximum of N excited electrons. These

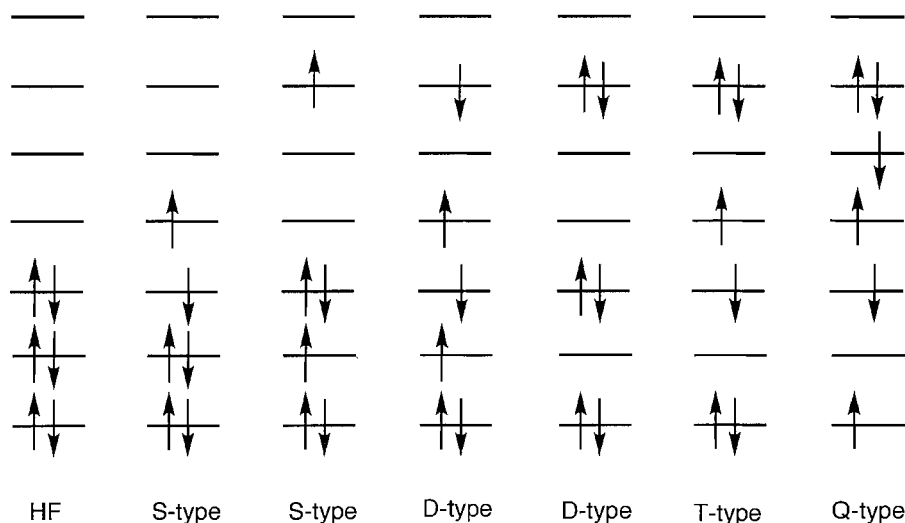


Figure 4.1 Excited Slater determinants generated from a HF reference

determinants are often referred to as *Singles (S)*, *Doubles (D)*, *Triples (T)*, *Quadruples (Q)* etc.

The total number of determinants that can be generated depends on the size of the basis set, the larger the basis, the more virtual MOs, and the more excited determinants can be constructed. If all possible determinants in a given basis set are included, all the electron correlation (in the given basis) is (or can be) recovered. For an infinite basis the Schrödinger equation is then solved exactly. Note that “exact” in this context is not the same as the experimental value, as the nuclei are assumed to have infinite masses (Born–Oppenheimer approximation) and relativistic effects are neglected. Methods which include electron correlation are thus two-dimensional, the larger the one-electron expansion (basis set size) and the larger the many-electron expansion (number of determinants), the better are the results. This is illustrated in Figure 4.2.

In order to calculate total energies with a “chemical accuracy” of ~ 1 kcal/mol, it is necessary to use sophisticated methods for including electron correlation and large basis sets, which is only computationally feasible for small systems. Instead the focus is usually on calculating relative energies, trying to make the errors as constant as possible.

EC \	Basic	Minimum	DZ	DZP	Infinite
HF (0%)						HF limit
10%						
.....						
100%						“Exact”

Figure 4.2 Convergence to the exact solution

The important chemical changes take place in the valence orbitals; the core orbitals are almost constant. In many cases the interest is therefore only in calculating the correlation energy associated with the valence electrons. Limiting the number of determinants to only those which can be generated by exciting the valence electrons is known as the *frozen core* approximation. In some cases the highest virtual orbitals corresponding to the anti-bonding combinations of the core orbitals are also removed from the correlation treatment (*frozen virtuals*). The frozen core approximation is not justified in terms of total energy; the correlation of the core electrons gives a substantial energy contribution. However, it is essentially a constant factor, which drops out when calculating relative energies. Furthermore, if we really want to calculate the core electron correlation, the standard basis sets are insufficient. In order to represent the angular correlation, higher angular momentum functions with the same radial size as the filled orbitals are needed, e.g. p- and d-functions with large exponents for correlating the 1s-electrons, as discussed in Section 5.4.5. Just allowing excitations of the core electrons in a standard basis set does not “correlate” the core electrons.

There are three main methods for calculating electron correlation: *Configuration Interaction (CI)*, *Many Body Perturbation Theory (MBPT)* and *Coupled Cluster (CC)*. A word of caution before we describe these methods in more details. The Slater determinants are composed of spin-MOs, but since the Hamilton operator is independent of spin, the spin dependence can be factored out. Furthermore, to facilitate notation, it is often assumed that the HF determinant is of the RHF type. Finally, many of the expressions below involve double summations over identical sets of functions. To ensure only the unique terms are included, one of the summation indices must be restricted. Alternatively, both indices can be allowed to run over all values, and the overcounting corrected by a factor of 1/2. Various combinations of these assumptions result in final expressions which differ by factors of 1/2, 1/4 etc. from those given here. In the present book the MOs are always spin-MOs, and conversion of a restricted summation to an unrestricted is always noted explicitly.

Finally a comment on notation. The quality of a calculation is given by the level of theory (i.e. how much electron correlation is included) and the size of the basis set. In a commonly used /-notation, introduced by J. A. Pople, this is denoted as “level/basis”. If nothing further is specified, this implies that the geometry is optimized at this level of theory. As discussed in Section 5.5, the geometry is usually much less sensitive to the theoretical level than relative energies, and high-level calculations are therefore often carried out using geometries optimized at a lower level. This is denoted as “level2/basis2//level1/basis1”, where the notation after the // indicates the level at which the geometry is optimized.

4.2 Configuration Interaction

This is perhaps the easiest method to understand. It is based on the variational principle (Appendix B), analogous to the HF method. The trial wave function is written as a linear combination of determinants with the expansion coefficients determined by requiring that the energy should be a minimum (or at least stationary), a procedure known as *Configuration Interaction (CI)*. The MOs used for building the excited Slater determinants are taken from a Hartree–Fock calculation and held fixed. Subscripts S, D, T etc. indicate determinants which are singly, doubly, triply etc. excited relative to the

HF configuration.

$$\Psi_{\text{CI}} = a_0 \Phi_{\text{SCF}} + \sum_{\text{S}} a_{\text{S}} \Phi_{\text{S}} + \sum_{\text{D}} a_{\text{D}} \Phi_{\text{D}} + \sum_{\text{T}} a_{\text{T}} \Phi_{\text{T}} \dots = \sum_{i=0} a_i \Phi_i \quad (4.2)$$

This is an example of a constrained optimization, the energy should be minimized under the constraint that the total CI wave function is normalized. Introducing a Lagrange multiplier (Section 14.6), this can be written as

$$L = \langle \Psi_{\text{CI}} | \mathbf{H} | \Psi_{\text{CI}} \rangle - \lambda [\langle \Psi_{\text{CI}} | \Psi_{\text{CI}} \rangle - 1] \quad (4.3)$$

The first bracket is the energy of the CI wave function, the second bracket is the norm of the wave function. In terms of determinants (eq. (4.2)), these can be written as

$$\begin{aligned} \langle \Psi_{\text{CI}} | \mathbf{H} | \Psi_{\text{CI}} \rangle &= \sum_{i=0} \sum_{j=0} a_i a_j \langle \Phi_i | \mathbf{H} | \Phi_j \rangle = \sum_{i=0} a_i^2 E_i + \sum_{i=0} \sum_{j \neq i} a_i a_j \langle \Phi_i | \mathbf{H} | \Phi_j \rangle \\ \langle \Psi_{\text{CI}} | \Psi_{\text{CI}} \rangle &= \sum_{i=0} \sum_{j=0} a_i a_j \langle \Phi_i | \Phi_j \rangle = \sum_{i=0} a_i^2 \langle \Phi_i | \Phi_i \rangle = \sum_{i=0} a_i^2 \end{aligned} \quad (4.4)$$

The diagonal elements in the sum involving the Hamilton operator are energies of the corresponding determinants. The overlap elements between different determinants are zero as they are built from orthogonal MOs (eq. (3.20)). The variational procedure corresponds to setting all the derivatives of the Lagrange function (4.3) with respect to the a_i expansion coefficients equal to zero.

$$\begin{aligned} \frac{\partial L}{\partial a_i} &= 2 \sum_j a_j \langle \Phi_i | \mathbf{H} | \Phi_j \rangle - 2\lambda a_i = 0 \\ a_i (\langle \Phi_i | \mathbf{H} | \Phi_i \rangle - \lambda) + \sum_{j \neq i} a_j \langle \Phi_i | \mathbf{H} | \Phi_j \rangle &= 0 \\ a_i (E_i - \lambda) + \sum_{j \neq i} a_j \langle \Phi_i | \mathbf{H} | \Phi_j \rangle &= 0 \end{aligned} \quad (4.5)$$

If there is only one determinant in the expansion ($a_0 = 1$), the last equation shows that the Lagrange multiplier is the (CI) energy, $\lambda = E$.

As there is one equation (4.5) for each i , the variational problem is transformed into solving a set of CI *secular equations*. Introducing the notation $H_{ij} = \langle \Phi_i | \mathbf{H} | \Phi_j \rangle$ the matrix equation becomes

$$\begin{pmatrix} H_{00} - E & H_{01} & \dots & H_{0j} & \dots \\ H_{10} & H_{11} - E & \dots & H_{1j} & \dots \\ \dots & \dots & \dots & \dots & \dots \\ H_{j0} & \dots & \dots & H_{jj} - E & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_j \\ \dots \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \dots \\ 0 \\ \dots \end{pmatrix} \quad (4.6)$$

which in shorthand notation may be written as $(\mathbf{H} - E\mathbf{I}) \mathbf{a} = \mathbf{0}$ or as $\mathbf{H}\mathbf{a} = E\mathbf{a}$. Solving the secular equations is equivalent to diagonalizing the CI matrix, see Chapter 13. The CI energy is obtained as the lowest eigenvalue of the CI matrix, and the corresponding eigenvector contains the a_i coefficients in front of the determinants in eq. (4.2). The second lowest eigenvalue corresponds to the first excited state etc.

4.5 Size Consistency and Size Extensivity

As mentioned above, full CI is impossible, except for very small systems. The only general applicable method is CISD. Consider now a series of CISD calculations in order to construct the interaction potential between two H_2 molecules as a function of the distance between them. Relative to the HF wave function, there will be determinants which correspond to single excitations on only one of the H_2 fragments (S-type determinants), single excitations on both (D-type determinants), and double excitations only on one of the H_2 fragments (also D-type determinants). This will be the case at all intermolecular distances, also when the separation is very large. In that case, however, the system is just two H_2 molecules, and we could consider calculating the energy instead as twice the energy of one H_2 molecule. A CISD calculation on one H_2 molecule would generate singly and doubly excited determinants, and multiplying this by two, would generate determinants which are quadruply excited for the combined H_4 system. A CISD calculation of two H_2 molecules separated by say 100 Å will not give the same energy as twice the results from a CISD calculation on one H_2 molecule (this will be lower). This problem is referred to a *Size Inconsistency*. A very similar, but not identical concept, is *Size Extensivity*. Size consistency is only defined if the two fragments are non-interacting (separated by say 100 Å) while size extensivity implies that the method scales properly with the number of particles, i.e. the fragments can be interacting (separated by say 5 Å). Full CI is size consistent (and extensive), but all forms of truncated CI are not. The lack of size extensivity is the reason why CISD recovers less and less electron correlation as the systems grow larger.

4.9 Coupled Cluster Methods

Perturbation methods add all types of corrections (S, D, T, Q etc.) to the reference wave function to a given order (2, 3, 4 etc.). The idea in *Coupled Cluster* (CC) methods is to include all corrections of a given type to infinite order.¹⁷ The (intermediate normalized) coupled cluster wave function is written as

$$\begin{aligned}\Psi_{cc} &= e^{\mathbf{T}}\Phi_0 \\ e^{\mathbf{T}} &= \mathbf{1} + \mathbf{T} + \frac{1}{2}\mathbf{T}^2 + \frac{1}{6}\mathbf{T}^3 + \dots = \sum_{k=0}^{\infty} \frac{1}{k!}\mathbf{T}^k\end{aligned}\quad (4.46)$$

where the cluster operator \mathbf{T} is given by

$$\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3 + \dots + \mathbf{T}_N \quad (4.47)$$

The \mathbf{T}_i operator acting on a HF reference wave function generates all i th excited Slater determinants.

$$\begin{aligned}\mathbf{T}_1\Phi_0 &= \sum_i^{\text{occ}} \sum_a^{\text{vir}} t_i^a \Phi_i^a \\ \mathbf{T}_2\Phi_0 &= \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} t_{ij}^{ab} \Phi_{ij}^{ab}\end{aligned}\quad (4.48)$$

It is customary to use the term *amplitudes* for the expansion coefficients t , which are equivalent to the a_i coefficients in eq. (4.1).

From eqs. (4.46) and (4.47) the exponential operator may be written as

$$\begin{aligned}e^{\mathbf{T}} &= \mathbf{1} + \mathbf{T}_1 + \left(\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2\right) + \left(\mathbf{T}_3 + \mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3\right) \\ &+ \left(\mathbf{T}_4 + \mathbf{T}_3\mathbf{T}_1 + \frac{1}{2}\mathbf{T}_2^2 + \frac{1}{2}\mathbf{T}_2\mathbf{T}_1^2 + \frac{1}{24}\mathbf{T}_1^4\right) + \dots\end{aligned}\quad (4.49)$$

The first term generates the reference HF and the second all singly excited states. The first parenthesis generates all doubly excited states, which may be considered as *connected* (\mathbf{T}_2) or *disconnected* (\mathbf{T}_1^2). The second parenthesis generates all triply excited states, which again may be either “true” (\mathbf{T}_3) or “product” triples ($\mathbf{T}_2\mathbf{T}_1$, \mathbf{T}_1^3). The quadruply excited states can similarly be viewed as composed of five terms, a true quadruple and four product terms. Physically a connected type such as \mathbf{T}_4 corresponds to four electrons interacting simultaneously, while a disconnected term such as \mathbf{T}_2^2 corresponds to two non-interacting pairs of interacting electrons.

With the coupled cluster wave function (4.46) the Schrödinger equation becomes

$$\mathbf{H}e^{\mathbf{T}}\Phi_0 = Ee^{\mathbf{T}}\Phi_0 \quad (4.50)$$

Multiplying from the left by Φ_0^* and integrating gives

$$\begin{aligned}\langle\Phi_0|\mathbf{H}e^{\mathbf{T}}|\Phi_0\rangle &= E_{\text{cc}}\langle\Phi_0|e^{\mathbf{T}}\Phi_0\rangle \\ \langle\Phi_0|\mathbf{H}e^{\mathbf{T}}|\Phi_0\rangle &= E_{\text{cc}}\langle\Phi_0|(\mathbf{1} + \mathbf{T}_1 + \mathbf{T}_2 + \dots)\Phi_0\rangle \\ E_{\text{cc}} &= \langle\Phi_0|\mathbf{H}e^{\mathbf{T}}|\Phi_0\rangle\end{aligned}\quad (4.51)$$

Expanding out the exponential in eq. (4.46) and using the fact that the Hamilton operator contains only one- and two-electron operators (eq. (3.24)) we get

$$\begin{aligned}E_{\text{cc}} &= \langle\Phi_0|\mathbf{H}|(\mathbf{1} + \mathbf{T}_1 + \mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2)\Phi_0\rangle \\ E_{\text{cc}} &= \langle\Phi_0|\mathbf{H}|\Phi_0\rangle + \langle\Phi_0|\mathbf{H}|\mathbf{T}_1\Phi_0\rangle + \langle\Phi_0|\mathbf{H}|\mathbf{T}_2\Phi_0\rangle + \frac{1}{2}\langle\Phi_0|\mathbf{H}|\mathbf{T}_1^2\Phi_0\rangle \\ E_{\text{cc}} &= E_0 + \sum_i^{\text{occ}} \sum_a^{\text{vir}} t_i^a \langle\Phi_0|\mathbf{H}|\Phi_i^a\rangle + \sum_{i<j}^{\text{occ}} \sum_{a<b}^{\text{vir}} (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) \langle\Phi_0|\mathbf{H}|\Phi_{ij}^{ab}\rangle\end{aligned}\quad (4.52)$$

When using HF orbitals for constructing the Slater determinants, the first matrix elements are zero (Brillouins theorem) and the second matrix elements are just

two-electron integrals over MOs (eq. (4.7)).

$$E_{cc} = E_0 + \sum_{i < j}^{\text{occ}} \sum_{a < b}^{\text{vir}} (t_{ij}^{ab} + t_i^a t_j^b - t_i^b t_j^a) (\langle \phi_i \phi_j | \phi_a \phi_b \rangle - \langle \phi_i \phi_j | \phi_b \phi_a \rangle) \quad (4.53)$$

The coupled cluster correlation energy is therefore determined completely by the singles and doubles amplitudes and the two-electron MO integrals.

Only equation for the amplitudes is obtained by multiplying the Schrödinger equation (4.50) from the left by a singly excited determinant $(\Phi_m^e)^*$ and integrating.

$$\begin{aligned} \langle \Phi_m^e | \mathbf{H} e^{\mathbf{T}} | \Phi_0 \rangle &= E_{cc} \langle \Phi_m^e | e^{\mathbf{T}} | \Phi_0 \rangle \\ \langle \Phi_m^e | \mathbf{H} | (1 + \mathbf{T}_1 + \mathbf{T}_2 + \frac{1}{2} \mathbf{T}_1^2 + \mathbf{T}_3 + \mathbf{T}_1 \mathbf{T}_2 + \frac{1}{6} \mathbf{T}_1^3) \Phi_0 \rangle &= E_{cc} \langle \Phi_m^e | \mathbf{T}_1 \Phi_0 \rangle \\ \langle \Phi_m^e | \mathbf{H} | \Phi_0 \rangle + \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_1 \Phi_0 \rangle + \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_2 \Phi_0 \rangle + \frac{1}{2} \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_1^2 \Phi_0 \rangle \\ + \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_3 \Phi_0 \rangle + \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_1 \mathbf{T}_2 \Phi_0 \rangle + \frac{1}{6} \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_1^3 \Phi_0 \rangle &= E_{cc} \langle \Phi_m^e | \mathbf{T}_1 \Phi_0 \rangle \end{aligned} \quad (4.54)$$

Only the indicated terms survive in the expansion when the orthogonality of the Slater determinants and the nature of the Hamilton operator is considered. The first term in the last equation is again zero due to Brillouins theorem, and the remaining forms a coupled set of equations of all singles, doubles and triples amplitudes (substituting in the expression for the energy, eq. (4.52)). Other equations connecting amplitudes may be obtained by multiplying from the left by a double, triple etc. excited determinant and integrating.

4.9.1 Truncated Coupled Cluster Methods

So far everything is exact. If all cluster operators up to \mathbf{T}_N are included in \mathbf{T} , all possible excited determinants are generated and the coupled cluster wave function is equivalent to full CI. This is as already stated impossible for all but the smallest systems. The cluster operator must therefore be truncated at some excitation level. When the \mathbf{T} operator is truncated, some of the terms in the amplitude equations will become zero, and the amplitudes derived from these approximate equations will no longer be exact. The energy calculated from these approximate singles and doubles amplitudes (eq. (4.53)) will therefore also be approximate. How severe the approximation is depends on how many terms are included in \mathbf{T} . Including only the \mathbf{T}_1 operator does not give any improvement over HF, as matrix elements between the HF and singly excited states are zero. The lowest level of approximation is therefore $\mathbf{T} = \mathbf{T}_2$, referred to as *Coupled Cluster Doubles* (CCD) Compared to the number of doubles, there are relatively few singly excited states. Using $\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2$ gives the CCSD model which is only slightly more demanding than CCD, and yields a more complete model. Both CCD and CCSD involve a computational effort which scales as M^6 in the limit of a large basis set. The next higher level has $\mathbf{T} = \mathbf{T}_1 + \mathbf{T}_2 + \mathbf{T}_3$, giving the CCSDT model¹⁸ This involves a computational effort which scales as M^8 , and is more demanding than CISDT. It (and higher-order methods like CCSDTQ) can consequently

only be used for small systems, and CCSD is the only generally applicable coupled cluster method.

Let us look in a bit more detail at the CCSD method. In this case we have (eqs. (4.46) and (4.47)).

$$e^{\mathbf{T}_1 + \mathbf{T}_2} = \mathbf{1} + \mathbf{T}_1 + \left(\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2\right) + \left(\mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3\right) + \left(\frac{1}{2}\mathbf{T}_2^2 + \frac{1}{2}\mathbf{T}_2\mathbf{T}_1^2 + \frac{1}{24}\mathbf{T}_1^4\right) + \dots \quad (4.55)$$

The CCSD energy is given by the general CC equation (4.53), and amplitude equations are derived by multiplying (4.50) with a singly excited determinant and integrating (analogously to eq. (4.54)).

$$\begin{aligned} \langle \Phi_m^e | \mathbf{H} | (\mathbf{1} + \mathbf{T}_1 + (\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2) + (\mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3)) \Phi_0 \rangle &= E_{\text{CCSD}} \langle \Phi_m^e | \mathbf{T}_1 \Phi_0 \rangle \\ \langle \Phi_m^e | \mathbf{H} | \Phi_0 \rangle + \langle \Phi_m^e | \mathbf{H} | \mathbf{T}_1 \Phi_0 \rangle + \langle \Phi_m^e | \mathbf{H} | (\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2) \Phi_0 \rangle \\ &\quad + \langle \Phi_m^e | \mathbf{H} | (\mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3) \Phi_0 \rangle = E_{\text{CCSD}} \sum_{ia} t_i^a \langle \Phi_m^e | \Phi_i^a \rangle \\ \langle \Phi_m^e | \mathbf{H} | \Phi_0 \rangle + \sum_{ia} t_i^a \langle \Phi_m^e | \mathbf{H} | \Phi_i^a \rangle + \sum_{ijab} (t_{ij}^{ab} + t_i^a t_j^b - t_j^b t_i^a) \langle \Phi_m^e | \mathbf{H} | \Phi_{ij}^{ab} \rangle \\ &\quad + \sum_{ijkabc} (t_{ij}^{ab} t_k^c + \dots + t_i^a t_j^b t_k^c + \dots) \langle \Phi_m^e | \mathbf{H} | \Phi_{ijk}^{abc} \rangle = E_{\text{CCSD}} t_m^e \end{aligned} \quad (4.56)$$

The notation $(t_i^a t_j^b t_k^c + \dots)$ indicates that several other terms involving permutations of the indices are omitted. Multiplying eq. (4.50) with a doubly excited determinant gives

$$\begin{aligned} \langle \Phi_{mn}^{ef} | \mathbf{H} | (1 + \mathbf{T}_1 + (\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2) + (\mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3) \\ + (\frac{1}{2}\mathbf{T}_2^2 + \frac{1}{2}\mathbf{T}_2\mathbf{T}_1^2 + \frac{1}{24}\mathbf{T}_1^4)) \Phi_0 \rangle &= E_{\text{CCSD}} \langle \Phi_{mn}^{ef} | (\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2) \Phi_0 \rangle \\ \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_0 \rangle + \langle \Phi_{mn}^{ef} | \mathbf{H} | \mathbf{T}_1 \Phi_0 \rangle \\ &\quad + \langle \Phi_{mn}^{ef} | \mathbf{H} | (\mathbf{T}_2 + \frac{1}{2}\mathbf{T}_1^2) \Phi_0 \rangle + \langle \Phi_{mn}^{ef} | \mathbf{H} | (\mathbf{T}_2\mathbf{T}_1 + \frac{1}{6}\mathbf{T}_1^3) \Phi_0 \rangle \\ &\quad + \langle \Phi_{mn}^{ef} | \mathbf{H} | (\frac{1}{2}\mathbf{T}_2^2 + \frac{1}{2}\mathbf{T}_2\mathbf{T}_1^2 + \frac{1}{24}\mathbf{T}_1^4) \Phi_0 \rangle = E_{\text{CCSD}} \sum_{ijab} (t_{ij}^{ab} + t_i^a t_j^b - t_j^b t_i^a) \langle \Phi_{mn}^{ef} | \Phi_{ij}^{ab} \rangle \\ \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_0 \rangle + \sum_{ia} t_i^a \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_i^a \rangle + \sum_{ijab} (t_{ij}^{ab} + t_i^a t_j^b - t_j^b t_i^a) \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_{ij}^{ab} \rangle \\ &\quad + \sum_{ijkabc} (t_{ij}^{ab} t_k^c + \dots + t_i^a t_j^b t_k^c + \dots) \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_{ijk}^{abc} \rangle \\ &\quad + \sum_{ijklabcd} (t_{ij}^{ab} t_{kl}^{cd} + \dots + t_{ij}^{ab} t_k^c t_l^d + \dots + t_i^a t_j^b t_k^c t_l^d + \dots) \langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_{ijkl}^{abcd} \rangle \\ &= E_{\text{CCSD}} (t_{mn}^{ef} + t_m^e t_n^f - t_m^f t_n^e) \end{aligned} \quad (4.57)$$

The equations (4.56) and (4.57) involve matrix elements between singles and triples and between doubles and quadruples. However, since the Hamilton operator only contains

one- and two-electron operators, these are actually identical to matrix elements between the reference and a doubly excited state. Consider for example $\langle \Phi_m^e | \mathbf{H} | \Phi_{ijk}^{abc} \rangle$. Unless m equals either i , j or k , and e equals either a , b , or c , there will be one overlap integral between different MOs which makes the matrix element zero. If for example $m = k$ and $e = c$, then the MO integral over these indices factor out as 1, and the rest is equal to a matrix element $\langle \Phi_0 | \mathbf{H} | \Phi_{ij}^{ab} \rangle$. Similarly, the matrix element $\langle \Phi_{mn}^{ef} | \mathbf{H} | \Phi_{ijkl}^{abcd} \rangle$ between a doubly and a quadruply excited determinant is only non-zero if mn matches up with two of the $ijkl$ indices, and ef matches up with $abcd$. Again such non-zero matrix elements are equal to matrix elements between the reference and a doubly excited determinant, eq. (4.7).

All the matrix elements can be evaluated in terms of MO integrals, and when the expression for the energy (4.53) is substituted into (4.56) and (4.57), they form coupled non-linear equations for the singles and doubles amplitudes. The equations contain terms up to quartic in the amplitudes, e.g. $(t_i^a)^4$ (since \mathbf{H} contains one- and two-electron operators), and must be solved by iterative techniques. Once the amplitudes are known, the energy and wave function can be calculated. The important aspect in coupled cluster methods is that excitations of higher order than the truncation of the \mathbf{T} operator enter the amplitude equation. Quadruply excited states, for example, are generated by the \mathbf{T}_2^2 operator in CCSD, and they enter the amplitude equations with a weight given as a product of doubles amplitudes. Quadruply excited states influence the doubles amplitudes, and thereby also the CCSD energy. It is the inclusion of these products of excitations that makes coupled cluster theory size extensive.

In the above the coupled cluster equations have been derived by multiplying the Schrödinger equation with $\langle \Phi_0 |$, $\langle \Phi_m^e |$ and $\langle \Phi_{mn}^{ef} |$. An alternative way of deriving the coupled cluster equations is to multiply with $\langle \Phi_0 | e^{-\mathbf{T}}$, $\langle \Phi_m^e | e^{-\mathbf{T}}$ and $\langle \Phi_{mn}^{ef} | e^{-\mathbf{T}}$. Just as $e^{\mathbf{T}}$ is an excitation operator working on the function to the right, $e^{-\mathbf{T}}$ is a de-excitation operator working on the function to the left. Thus $\langle \Phi_0 | e^{-\mathbf{T}}$ tries to generate de-excitations from the reference, which is impossible, i.e. $\langle \Phi_0 | e^{-\mathbf{T}} = \langle \Phi_0 |$. However $\langle \Phi_m^e | e^{-\mathbf{T}}$ generates in addition to the singly excited state also the reference wave function. Similarly $\langle \Phi_{mn}^{ef} | e^{-\mathbf{T}}$ generates additionally the reference and singly excited states. The main advantage is that the coupled cluster equations are obtained directly without having to substitute in the expression for the energy in the amplitude equations.

Overview of Gaussian Input

In This Appendix:

**Gaussian Input
File Format**

**Constructing
Z-matrices**

We've already looked briefly at *Gaussian* input in the *Quick Start*. Here we present a more formal definition and discuss the various molecule specification options.

Gaussian input is designed to be free-format and extremely flexible. For example, it is not case-sensitive, and keywords and options may be shortened to a unique abbreviation.

Input File Sections

Gaussian input (which is the same for all versions of the program, including the Window version) has the basic structure described in the following table. Note that the input sections marked with an asterisk are required in every input file:

Input Section	Contents
Link 0 Commands	Defines the locations of scratch files and job resource limits.
*Route Section	Specifies the job type and model chemistry.
*blank line	Separates the route section from the title section.
*Title Section	Describes the job for the output and archive entry.
*blank line	
*Molecule Specification	Gives the structure of the molecule to be studied.
*blank line	
Variables Section	Specifies values for the variables used in the molecule specification.
blank line	

Note that the separate input sections are separated from one another by blank lines. These blank lines are inserted automatically into input files created with the **Job Entry** window in the Windows version and need not be entered by the user. If you choose to create a *Gaussian* input file using an external Windows editor, however, you must follow the same rules for input as under other versions of *Gaussian*.

Note that some job types require additional sections not listed.

Input lines have a maximum length of 80 characters.



The Route Section

The first line of the route section always begins with a pound sign (#) in the first column. This section specifies the theoretical procedure, basis set, and desired type of calculation. It may also include other keywords. The ordering of keywords is not important. Some keywords require options; the following input line illustrates the possible formats for keywords with options:

```
#T RHF/6-31G(d) SCF=Tight      Units=(Bohr,Radian)   Opt Test
Keyword with:      a single option      ≥ 2 options          no options
```

The amount of spacing between items is not significant in *Gaussian* input. In the route section, commas or slashes may be substituted for spaces if desired (except within parenthesized options, where slashes don't work). For example, the previous route section used a slash to separate the procedure and basis set, spaces to separate other keywords, and commas to separate the options to the **Units** keyword.

The route section may extend over more than one line if necessary. Only the first line need begin with a pound sign, although any others may. The route section is terminated by a blank line.

The Title Section

The title section consists of one or more lines of descriptive information about the job. It is included in the output and in the archive entry but is not otherwise used by *Gaussian*. This section is terminated by a blank line.

Specifying Molecular Structures

Gaussian accepts molecule specifications in several different formats:

- ◆ Cartesian coordinates
- ◆ Z-matrix format (internal coordinates)
- ◆ Mixed internal and Cartesian coordinates

All molecule specifications require that the *charge* and *spin multiplicity* be specified (as two integers) on the first line of this section. The charge is a positive or negative integer specifying the total charge on the molecule. Thus, 1 or +1 would be used for a singly-charged cation, -1 designates a singly-charged anion, and 0 represents a neutral molecule.

Spin Multiplicity

The *spin multiplicity* is given by the equation $2S + 1$, where S is the total spin for the molecule. Paired electrons contribute nothing to this quantity. They have a net spin of zero since an alpha electron has a spin of $+\frac{1}{2}$ and a beta electron has a spin of $-\frac{1}{2}$. Each unpaired electron contributes $+\frac{1}{2}$ to S . Thus, a singlet—a system with no unpaired electrons—has a spin multiplicity of 1, a doublet (one unpaired electron)

has a spin multiplicity of 2, a triplet (two unpaired electrons of like spin) has a spin multiplicity of 3, and so on.

Units

The units in a Z-matrix are angstroms for lengths and degrees for angles by default; the default units for Cartesian coordinates are angstroms. These are also the default units for lengths and angles used in *Gaussian* output. You can change them to bohrs and/or radians by including the **Units** keyword in the route section with one or both of its options: **Bohr** and **Radian**.

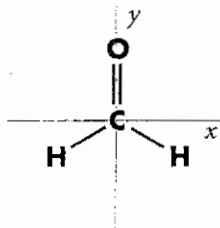
Cartesian Coordinate Input

Cartesian coordinate input consists of a series of lines of the form:

Atomic-symbol X-coordinate Y-coordinate Z-coordinate

For example, here is the molecular structure for formaldehyde, given in Cartesian coordinates:

```
O 1
C 0.0 0.0 0.0
O 0.0 1.22 0.0
H 0.94 -0.54 0.0
H -0.94 -0.54 0.0
```



Z-Matrix Input

The other syntax for supplying molecular structures to *Gaussian 94* is the *Z-matrix*. A Z-matrix specifies the locations of and bonds between atoms using bond lengths, bond angles, and dihedral (torsion) angles.

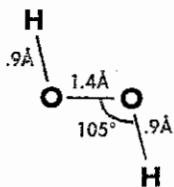
Each atom in the molecule is described on a separate input line within the Z-matrix. As we consider the procedure for creating a Z-matrix, we'll use hydrogen peroxide as an example. These are the steps to do so:

1. **Choose a starting atom in the molecule, and conceptually place it at the origin in three dimensional space.**

The first line of the Z-matrix consists solely of the label for this atom. An atom label is made up of its atomic symbol optionally followed by an integer (no spaces), used to distinguish it from the other atoms of the same type (e.g. H1 for the first hydrogen, H2 for the second one, and so on).

We'll use the left oxygen atom in the illustration as our first atom:

```
O1
```



2. Choose another atom bonded to the first atom. Place it along the Z-axis, and specify the length of the bond connecting the two atoms.

This second input line will include the atom label of the second atom, the label of the atom it is bonded to (the first atom), and the bond length, in that order. Items may be separated by spaces, tabs or commas.

We'll use the hydrogen atom bonded to the first oxygen for our second atom:

```
O1
H1 O1 .9
```

3. Choose a third atom bonded to either of the previous two atoms and specify the angle formed by the two bonds.

This angle locates the molecule's position in the XZ-plane. This input line will include the new atom's label, the atom it is bonded to and the bond length, the label of the other atom forming the bond angle, and the angle's value.

We have only one choice for the third atom in our Z-matrix, the second oxygen atom:

```
O1
H1 O1 0.9
O2 O1 1.4 H1 105.0
```

The new line illustrates an important point about numeric values within Z-matrices. Since they correspond to floating point quantities, they must include a decimal point, as in the value above. This is true even for a value of 0.

4. Describe the positions of all subsequent atoms by specifying:

- ◆ Its atom label.
- ◆ An atom it is bonded to and the bond length.
- ◆ A third atom bonded to it (or to the second atom), and the value of the resulting bond angle.
- ◆ A fourth atom bonded to either end of the previous chain, and the value of the dihedral (torsion) angle formed by the four atoms.

Dihedral angles describe the angle the fourth atom makes with respect to the plane defined by the first three atoms; their values range from 0 to 360 degrees, or from -180 to 180 degrees. Dihedral angles are easy to visualize using Newman projections. The illustration shows the Newman projection for hydrogen peroxide, looking down the O-O bond. Positive dihedral angles correspond to clockwise rotation in the Newman projection.



Obviously, we'll use the remaining hydrogen atom for the fourth line of our Z-matrix for hydrogen peroxide. Here is the completed molecule specification:

```

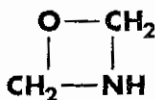
0 1                               Charge and multiplicity.
O1                                Oxygen atom #1.
H1 O1 .9                          Hydrogen #1, connected to oxygen #1 by a bond of 0.9 Å.
O2 O1 1.4 H1 105.                 Oxygen #2: O2-O1 = 1.4 Å; -H1-O1-O2 = 105°.
H2 O2 .9 O1 105. H1 120.         Hydrogen #2: H2-O2 bond = 0.9 Å; -H2-O2-O1 = 105°;
                                   dihedral angle H2-O2-O1-H1 = 120°.

```

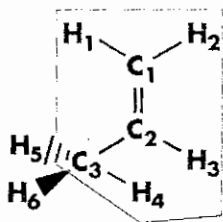
Sources for bond lengths, bond angles, and dihedral angles include the published literature, standard references like the CRC series, and previous calculations. Z-matrices may also be created by the **NewZMat** utility from data generated by a wide variety of drawing packages. Refer to the *Quick Start* for a sample conversion operation for your version of *Gaussian*.

Mixed Internal and Cartesian Coordinates

It is also possible to specify the molecular structure in a format which combines Cartesian coordinates and Z-matrix style input; this format is referred to as mixed internal and Cartesian coordinates. It is useful for systems where some parts of the molecule are more easily specified in Cartesian coordinates and others are more easily described as a Z-matrix. Consult Exercise C.2 (page 293) and Appendix B of the *Gaussian 94 User's Reference* for more information on this topic.



More Complex Z-Matrices



Constructing a Z-matrix for propene provides a more challenging example. Note that the atoms inside the red line in the illustration all lie in a plane. First we'll specify the carbon atoms:

```

0 1                               Charge and multiplicity.
C1                                Carbon atom at the head of the double bond.
C2 C1 1.34                         Carbon atom on the other end of the double bond.
C3 C2 1.52 C1 120.                 Third carbon atom.

```

Next, we'll specify the hydrogens on C1 and C2. The bond angles formed with the double-bonded carbons and each of these hydrogens is 120° . We'll pick simple dihedral angles for each of them: