

Ringo

A HeH⁺ RHF/STO-3G Calculation Program

St Maxwell

(st_maxwell@outlook.com)

1 前言

这是我写的第一个量子化学程序，功能很简单，就是 HeH⁺ 的 RHF/STO-3G 计算。首先说说程序的名字「Ringo」的由来。好吧，其实只是因为我喜欢椎名林檎，所以就起了 Ringo 这个名字。然后我再考虑了一下，觉得可以这样理解：「Ringo Is Not Gaussian/Orca」。顺便还能中二一下，暗指以后可以与这两个软件平起平坐。当然现在说这个很不现实啦。

从今年年初的时候开始看 Levine 的 *Quantum Chemistry*，到年中时候看 *Modern Quantum Chemistry*，再到现在写出第一个量化程序，我想这一年我还是收获了很多。虽然现在对 Hartree Fock 方法有一定初步的了解，但无论从理论还是编程的方面来讲，我都只是初学者。写这份文档的目的，很大程度并不是为了去教别人如何编写量化程序，而是作为学习的一个过程。从看书，做笔记，写程序，写文档这一路下来，我觉得比单纯完成其中的一部分能够收获更多。举个例子，看书抄笔记的时候，有时候觉得书上说什么都挺有道理的，而在写程序的时候，才会真正的咀嚼书中的内容。

当然啦，如果你也是正在学习理论化学的一员，我的代码和这份文档对你能有些许帮助的话，我也会非常开心的。

2 Hartree Fock 方法

这里并不会完整地介绍 Hartree Fock 方法，大部分时候只是把公式列出来。电子积分等计算的公式，大多都是 *Modern Quantum Chemistry* 上介绍的，并没有很高端的算法。

2.1 Roothaan 方程

Roothaan 方程具有本征方程的形式

$$f(\mathbf{r}_1)\psi_i(\mathbf{r}_1) = \varepsilon_i\psi_i(\mathbf{r}_1) \quad (1)$$

式中 f 为 Fock 算符， $\psi_i(\mathbf{r}_1)$ 为空间轨道形式的分子轨道，可由 K 个已知的基函数线性展开：

$$\psi_i = \sum_{\nu=1}^K C_{\nu i} \phi_{\nu}, \quad \nu = 1, 2, \dots, K \quad (2)$$

也可用矩阵的形式表示：

$$(\psi_1, \psi_2, \dots, \psi_N) = (\phi_1, \phi_2, \dots, \phi_K) \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1N} \\ C_{21} & C_{22} & \dots & C_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ C_{K1} & C_{K2} & \dots & C_{KN} \end{pmatrix} \quad (3)$$

若我们用一组已知的基函数进行展开，那么问题就变成了求解一组展开系数。在(1)式两边左乘 ϕ_μ^* 并积分。

$$\sum_{\nu} C_{\nu i} \int d\mathbf{r}_1 \phi_\mu^*(1) f(1) \phi_\nu(1) = \varepsilon_i \sum_{\nu} C_{\nu i} \int d\mathbf{r}_1 \phi_\mu^*(1) \phi_\nu(1) \quad (4)$$

定义重叠矩阵 **S**

$$S_{\mu\nu} = \int d\mathbf{r}_1 \phi_\mu^*(1) \phi_\nu(1) \quad (5)$$

我们使用的基函数并不一定是正交归一的，所以重叠矩阵可能不是单位矩阵。但是可以通过酉阵进行对角化。

Fock 矩阵 **F**

$$F_{\mu\nu} = \int d\mathbf{r}_1 \phi_\mu^*(1) f(1) \phi_\nu(1) \quad (6)$$

所以可以将 Roothaan 方程改写为

$$\sum_{\nu} F_{\mu\nu} C_{\nu i} = \varepsilon_i \sum_{\nu} S_{\mu\nu} C_{\nu i} \quad (7)$$

同样可以有矩阵的形式

$$\mathbf{FC} = \mathbf{SC}\epsilon \quad (8)$$

注意上式中 ϵ 为对角阵。

2.2 电荷密度

$$\rho(\mathbf{r}) = 2 \sum_a^{N/2} |\psi_a(\mathbf{r})|^2 \quad (9)$$

上式中的 $N/2$ 是因为体系为闭壳层，更合适的表示应该是遍历所有占据轨道。将分子轨道的展开式(2)代入到上式中，结果是（可以自己推一下）：

$$\begin{aligned} \rho(\mathbf{r}) &= \sum_{\mu\nu} \left[2 \sum_a^{N/2} C_{\mu a} C_{\nu a} \right] \phi_\mu(\mathbf{r}) \phi_\nu^*(\mathbf{r}) \\ &= \sum_{\mu\nu} P_{\mu\nu} \phi_\mu(\mathbf{r}) \phi_\nu^*(\mathbf{r}) \end{aligned} \quad (10)$$

定义了密度矩阵，也称为电荷密度键级矩阵：

$$P_{\mu\nu} = 2 \sum_a^{N/2} C_{\mu a} C_{\nu a} \quad (11)$$

2.3 Fock 矩阵

$$\begin{aligned}
F_{\mu\nu} &= \int d\mathbf{r}_1 \phi_\mu^*(1) f(1) \phi_\nu(1) \\
&= \int d\mathbf{r}_1 \phi_\mu^*(1) h(1) \phi_\nu(1) + \sum_a^{N/2} \int d\mathbf{r}_1 \phi_\mu^* [2J_a(1) - K_a(1)] \phi_\nu(1) \\
&= h_{\mu\nu} + \sum_a^{N/2} [2(\mu\nu|aa) - (\mu a|a\nu)]
\end{aligned} \tag{12}$$

$h_{\mu\nu}$ 也可写成 $H_{\mu\nu}^{\text{core}}$, 称为核 Hamiltonian 矩阵, 其中的矩阵元包含了动能积分和核库仑积分。

$$T_{\mu\nu} = \int d\mathbf{r}_1 \phi_\mu^*(\mathbf{r}_1) \left[-\frac{1}{2} \nabla_1^2 \right] \phi_\nu(\mathbf{r}_1) \tag{13}$$

$$V_{\mu\nu}^{\text{nucl}} = \int d\mathbf{r}_1 \phi_\mu^*(\mathbf{r}_1) \left[-\sum_A \frac{Z_A}{|\mathbf{r}_1 - \mathbf{R}_A|} \right] \phi_\nu(\mathbf{r}_1) \tag{14}$$

$$H_{\mu\nu}^{\text{core}} = T_{\mu\nu} + V_{\mu\nu}^{\text{nucl}} \tag{15}$$

给定了基组 $\{\phi_\mu\}$ 后, 核 Hamiltonian 矩阵只需要计算一次, 之后不会再变化。对于 Fock 矩阵剩余的部分, 将分子轨道的展开式代入到双电子积分项。

$$\begin{aligned}
F_{\mu\nu} &= H_{\mu\nu}^{\text{core}} + \sum_a^{N/2} \sum_{\lambda\sigma} C_{\lambda a} C_{\sigma a}^* [2(\mu\nu|\lambda\sigma) - (\mu\lambda|\sigma\nu)] \\
&= H_{\mu\nu}^{\text{core}} + \sum_{\lambda\sigma} P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\lambda|\sigma\nu) \right] \\
&= H_{\mu\nu}^{\text{core}} + G_{\mu\nu}
\end{aligned} \tag{16}$$

对了, 双电子积分的公式为:

$$(\mu\nu|\lambda\sigma) = \int d\mathbf{r}_1 d\mathbf{r}_2 \phi_\mu^*(\mathbf{r}_1) \phi_\nu(\mathbf{r}_1) r_{12}^{-1} \phi_\lambda^*(\mathbf{r}_2) \phi_\sigma(\mathbf{r}_2) \tag{17}$$

实际计算中, 会利用对称性省去很多组双电子积分的计算。假如使用的基函数的数目为 N , 那么一共有 N^4 个双电子积分。对于我们的体系, 全部的双电子积分共有 16 个, 但实际只需要计算 6 个不同的即可。不过在程序中, 由于我暂时没有考虑如何利用对称性, 所以计算了全部的双电子积分。好在计算的体系很小, 而且程序是专门算这个的, 所以对计算速度的影响不大。随着体系的增大, 提高双电子积分的计算速度是很关键的一点。

2.4 基的正交化

重叠矩阵 \mathbf{S} 通常不是单位矩阵, 即我们使用的基组不是正交归一化的。但是我们可以通过变换矩阵将其对角化。

$$\mathbf{X}^\dagger \mathbf{S} \mathbf{X} = \mathbf{E}$$

这里介绍正则对角化。考虑本征值问题 $\mathbf{S} \mathbf{U} = \mathbf{U} \mathbf{s}$, 其中 \mathbf{U} 为本征矩阵, 是酉阵; \mathbf{s} 是 \mathbf{S} 的本征向量, 是对角阵, 其对角矩阵元为 \mathbf{S} 的本征值。通过酉阵 \mathbf{U} , 可以将重叠矩阵对角化:

$$\mathbf{U}^\dagger \mathbf{S} \mathbf{U} = \mathbf{s} \tag{18}$$

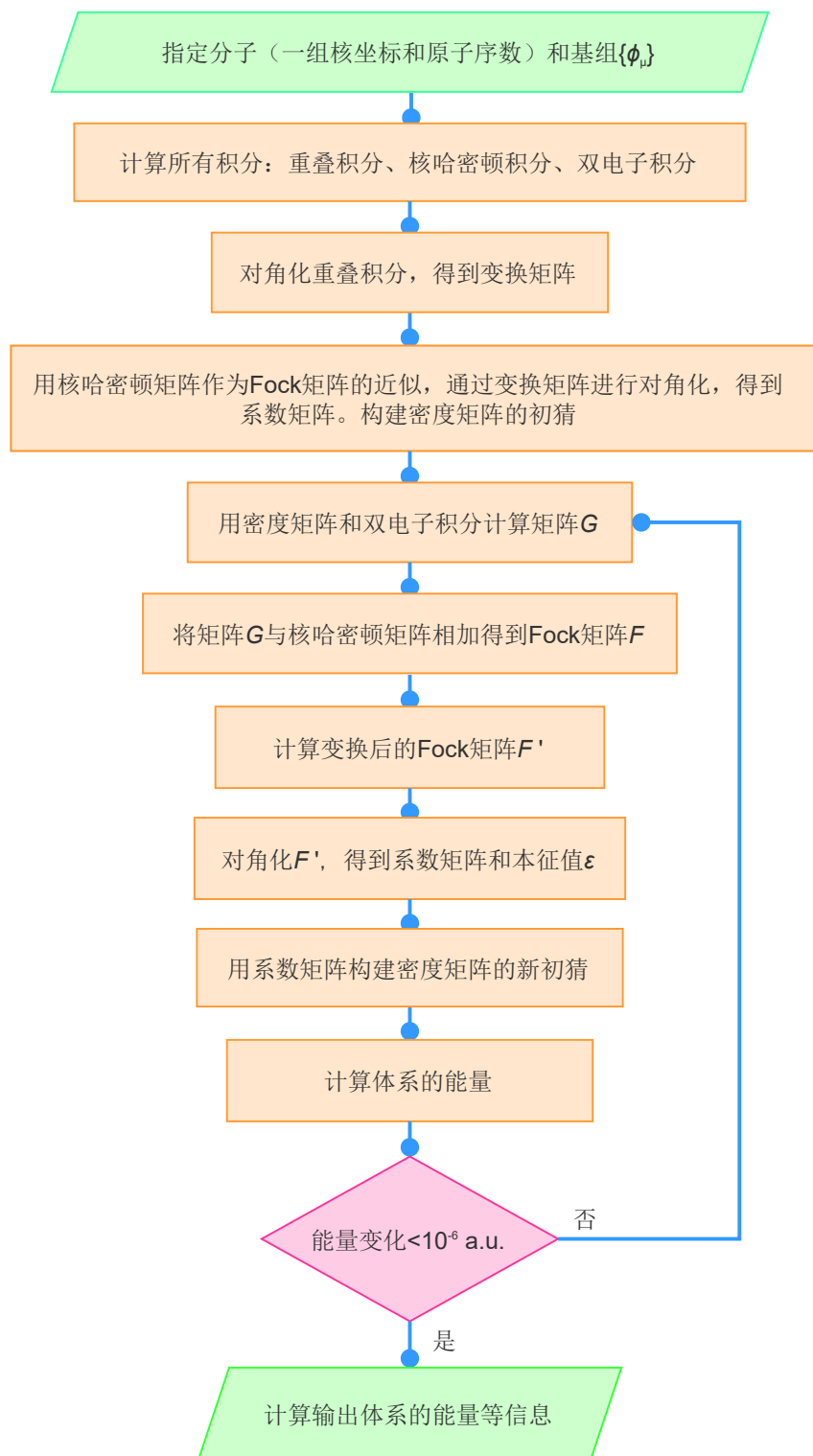


图 1. SCF 流程图

在两边同时乘以 $\mathbf{s}^{-1/2}$:

$$\begin{aligned}\mathbf{s}^{-1/2}\mathbf{U}^\dagger\mathbf{S}\mathbf{U}\mathbf{s}^{-1/2} &= \mathbf{s}^{-1/2}\mathbf{S}\mathbf{s}^{-1/2} \\ \mathbf{X}^\dagger\mathbf{S}\mathbf{X} &= \mathbf{E}\end{aligned}\tag{19}$$

得到变换矩阵 $\mathbf{X} = \mathbf{U}\mathbf{s}^{-1/2}$ 。

得到变换矩阵后，是否直接用它将基组进行正交归一化呢？通常并不会这样做，因为将基组变换后，我们需要重新计算所有的双电子积分，而计算双电子积分是 SCF 过程中最为耗时的。

一种效率更高的方式：

考虑一个新的系数矩阵 \mathbf{C}' ,

$$\mathbf{C}' = \mathbf{X}^{-1}\mathbf{C}, \quad \mathbf{C} = \mathbf{X}\mathbf{C}'$$

将上式代入到 Roothaan 方程：

$$\mathbf{F}\mathbf{X}\mathbf{C}' = \mathbf{S}\mathbf{X}\mathbf{C}'\varepsilon\tag{20}$$

左乘 \mathbf{X}^\dagger

$$(\mathbf{X}^\dagger\mathbf{F}\mathbf{X})\mathbf{C}' = (\mathbf{X}^\dagger\mathbf{S}\mathbf{X})\mathbf{C}'\varepsilon\tag{21}$$

定义新矩阵 \mathbf{F}'

$$\mathbf{F}' = \mathbf{X}^\dagger\mathbf{F}\mathbf{X}\tag{22}$$

$$\mathbf{F}'\mathbf{C}' = \mathbf{C}'\varepsilon\tag{23}$$

通过对角化 \mathbf{F}' ，可得到本征值 ε ，也可求得 \mathbf{C}' ，进而求出 \mathbf{C} 。

3 SCF

SCF 过程的流程可以看图(1)。

这里提几个点。我们计算时常用的基组为 Gaussian 型轨道，例如将用的 STO-3G 基组。每个基函数 ϕ_μ 是三个 Gaussian 函数的线性组合；密度矩阵的第一个初猜就是全为 0；收敛限只考察能量（也可以选择密度矩阵，不过计算各项总的标准差相对麻烦），收敛限为 $\delta\varepsilon < 10^{-6}$ 。

3.1 STO-3G

STO-3G 基组是用三个 Gaussian 函数（这三个函数被称为 Gaussian 原函数）拟合一个 Slater 函数（类氢原子的径向分布函数）。

$$\phi_{1s}^{\text{CGF}} = d_1\phi_{1s}^{\text{GF}}(\alpha_1) + d_2\phi_{1s}^{\text{GF}}(\alpha_2) + d_3\phi_{1s}^{\text{GF}}(\alpha_3)\tag{24}$$

$$\phi_{1s}^{\text{GF}}(\alpha, \mathbf{r} - \mathbf{R}_A) = (2\alpha/\pi)^{3/4} \exp(-\alpha|\mathbf{r} - \mathbf{R}_A|^2)\tag{25}$$

(24)式中的 d 和 α 分别为收缩系数和收缩指数，这两个值都是事先知道的（可通过与实验值拟合等方法确定），在 SCF 过程中不再变化。在计算中，基函数的积分被转化为 Gaussian 原函数之间的积分与求和。我们使用的电子积分的公式就是基于 Gaussian 函数推导出的。

在 Gaussian 软件中，可使用 `HF/STO-3G GFInput` 关键词在输出文件中打印出各原子的基组信息。我们用 HeH^+ 为例子，第一段是氢的基组信息，第二段是氦的。第一列是收缩指数 α ，第二列是收缩系数 d 。你可以使用这里的数据进行计算，不过我们是使用 *Modern Quantum Chemistry* 上提供的值。在后面会展示不同基组定义带来的结果的差异。

```
A0 basis set in the form of general basis input (Overlap normalization):
```

```
1 0
S 3 1.00      0.0000000000000
   0.3425250914D+01  0.1543289673D+00
   0.6239137298D+00  0.5353281423D+00
   0.1688554040D+00  0.4446345422D+00
****
2 0
S 3 1.00      0.0000000000000
   0.6362421394D+01  0.1543289673D+00
   0.1158922999D+01  0.5353281423D+00
   0.3136497915D+00  0.4446345422D+00
****
```

4 电子积分

我们使用的电子积分的公式主要来自 *Modern Quantum Chemistry* 的附录 A，这些公式只适用于 $1s$ 轨道波函数。其实比较推荐直接看书，毕竟上面有推导，而且公式中的符号都有说明。再提醒一点，这书附录中的公式是没有考虑 Gaussian 函数前面的归一化系数的，用的话需要自行加上。这里我是把归一化系数加上的。

4.1 重叠积分

$$(A|B) = \frac{(4\alpha\beta)^{3/4}}{(\alpha + \beta)^{3/2}} \cdot \exp\left(-\frac{\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2\right) \quad (26)$$

4.2 动能积分

$$(A|-\frac{1}{2}\nabla^2|B) = \frac{2^{3/2}(\alpha\beta)^{7/4}}{(\alpha + \beta)^{5/2}} \cdot \left(3 - \frac{2\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2\right) \cdot \exp\left(-\frac{\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2\right) \quad (27)$$

4.3 核库仑积分

$$(A|-\frac{Z_C}{r_{1C}}|B) = -\frac{(4\alpha\beta)^{4/3}}{\alpha + \beta}Z_C \cdot \exp\left(-\frac{\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2\right) \cdot \frac{\text{erf}(t^{1/2})}{t^{1/2}} \quad (28)$$

上式中 $t = (\alpha + \beta)|\mathbf{R}_P - \mathbf{R}_C|^2$ ，中心为 A 的 Gaussian 函数和中心为 B 的 Gaussian 函数相乘可等效为一个中心为 P 的 Gaussian 函数， $\mathbf{R}_P = (\alpha\mathbf{R}_A + \beta\mathbf{R}_B)/(\alpha + \beta)$ 。 erf 为误差函数，在 IVF 中是内置的函数。

这里还有一个不得不提到的坑，就是上面这个公式不适用于 A 、 B 、 C 在同一点的情况。当 $\mathbf{R}_A = \mathbf{R}_B = \mathbf{R}_C$ 的时候，上式中 $\mathbf{R}_P = \mathbf{R}_A$ ，所以 $t = 0$ 。会发现公式中出现了除以 0 的情况。我搜索之后，在 *Ideas of Quantum Chemistry* 的附录 P 中找到了公式。

$$(A| - Z_C/r_{1C}|B) = -Z_C \frac{2^{5/2}}{\pi^{1/2}} \frac{(\alpha\beta)^{3/4}}{\alpha + \beta} \quad (29)$$

所以在写程序时，注意这里应当加一个判断语句。

4.4 双电子积分

$$(AB|CD) = \frac{8(\alpha\beta\delta\gamma)^{3/4}}{(\alpha + \beta)(\delta + \gamma)(\alpha + \beta + \delta + \gamma)^{1/2}} \cdot \exp\left(-\frac{\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2 - \frac{\delta\gamma}{\delta + \gamma}|\mathbf{R}_C - \mathbf{R}_D|^2\right) \cdot \frac{\text{erf}(t^{1/2})}{t^{1/2}} \quad (30)$$

这里的 $t = (\alpha + \beta)(\delta + \gamma)/(\alpha + \beta + \delta + \gamma) \cdot |\mathbf{R}_Q - \mathbf{Q}_D|^2$ 。至于 \mathbf{R}_P 和 \mathbf{R}_Q 长什么样，我就不必再写一次了。同样地，当 $\mathbf{R}_P = \mathbf{R}_Q$ 的时候，就需要换另外一个公式：

$$(AB|CD) = \frac{16(\alpha\beta\delta\gamma)^{3/4}}{(\alpha + \beta)(\delta + \gamma)(\alpha + \beta + \delta + \gamma)^{1/2}} \cdot \exp\left(-\frac{\alpha\beta}{\alpha + \beta}|\mathbf{R}_A - \mathbf{R}_B|^2 - \frac{\delta\gamma}{\delta + \gamma}|\mathbf{R}_C - \mathbf{R}_D|^2\right) \quad (31)$$

5 计算

我们计算的体系为 HeH^+ ，取的键长为 1.4632 a.u.。氢原子的标号为 1，氦原子为 2。基组参数如下（我模仿了 *Gaussian* 的打印格式）：

```
AO basis set in the form of general basis input (Overlap normalization):
H      0
S      3 1.00      0.00000000000000
        3.4252509140D+00  0.1543289673D+00
        0.6239137298D+00  0.5353281423D+00
        0.1688554040D+00  0.4446345422D+00
****
He      0
S      3 1.00      0.00000000000000
        9.753934616D+00  0.1543289673D+00
        1.776691148D+00  0.5353281423D+00
        0.480844290D+00  0.4446345422D+00
****
```

5.1 结果

具体每一步得到的矩阵可以参考 *Modern Quantum Chemistry*, 这里我们只展示程序运行得到的结果。

```
Ringo: a HeH+ RHF/STO-3G calculation program
Ringo Is Not Gaussian/Orca
Programmed by St Maxwell
*****

Calulation Done Successfully!

Distance = 1.4632 a.u.

----- Restricted Hartree Fock Energy -----
Electronic energy = -4.227526 a.u.
Nuclear repulsion energy = 1.366867 a.u.
Total energy = -2.860659 a.u.

----- Population Analysis -----
Occupied orbital eigenvalue = -1.597452 a.u.
Virtual orbital eigenvalue = -0.061670 a.u.

----- Molecular Orbitals Coefficients -----
                phi1      phi2
Occupied orbital -- -0.80192 -0.33680
Virtual orbital -- -0.78226  1.06844

----- Mulliken Charges -----
He  0.470365
H   0.529635

Job cpu time: 0.0000 seconds.
```

5.2 对结果的说明

Electronic energy 指体系总的电子能。

$$E_0 = \frac{1}{2} \sum_{\mu\nu} P_{\mu\nu} (H_{\mu\nu}^{\text{core}} + F_{\mu\nu}) \quad (32)$$

Nuclear repulsion energy 是指两个原子核之间的库仑斥力：

$$E_{\text{repulsion}} = \frac{Z_{\text{H}}Z_{\text{He}}}{R_{\text{bond}}} \quad (33)$$

将电子能与核间排斥能即为体系的总能量，通常在量化程序中计算出的单点能就是这个。

布居分析中占据轨道和虚轨道的本征值是指轨道能，是 Fock 矩阵的本征值。注意，体系的电子能并不等于各个轨道的轨道能的直接求和。这一点可以看《[结构化学的随意笔记和想法](#)》中「轨道能」一节。

分子轨道系数为系数矩阵 \mathbf{C} 的矩阵元。

最后一段为 Mulliken 电荷，之前在电荷密度那一节没有提到。体系中的电子数为 N ，有以下关系：

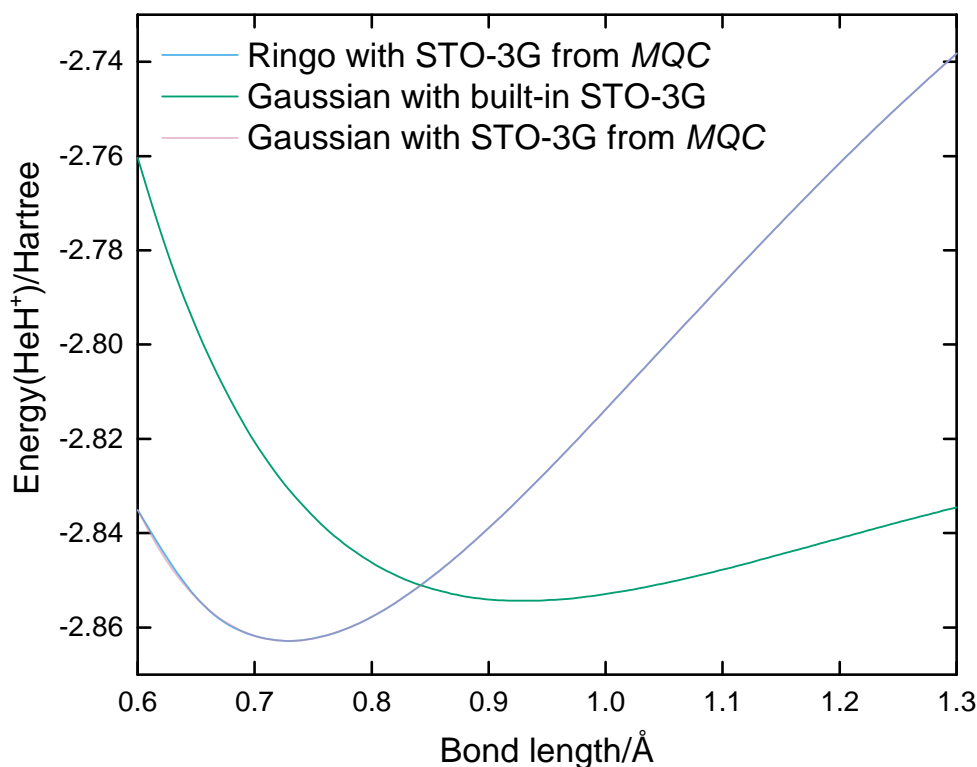
$$\begin{aligned} N &= 2 \sum_a^{\text{occupied}} \int d\mathbf{r} |\psi_a(\mathbf{r})|^2 \\ &= \sum_{\mu\nu} P_{\mu\nu} S_{\nu\mu} \\ &= \sum_{\mu} (\mathbf{PS})_{\mu\mu} \end{aligned} \quad (34)$$

分子中的某一原子上的净电荷为：

$$q_A = Z_A - \sum_{\mu \in A} (\mathbf{PS})_{\mu\mu} \quad (35)$$

$\mu \in A$ 的意思是指所有中心在 A 原子上的基函数。对于本例，氢原子对应的是记号为 1 的基函数， μ 的取值为 1。

之后用编写的程序考察了体系的总能量随键长的变化，并与 *Gaussian* 计算的结果相比较。可以看到，在使用相同定义的基组的情况下，两者的结果是一致的。



6 结语

限于我的理论水平和目前的精力，这份文档暂时只写到这里。随着我未来的学习和对程序的改进，我会逐渐为这个文档写下更多的内容。

7 参考资料

写这个程序的过程中，参考了以下资料：

1. *Modern Quantum Chemistry*。这是最重要的一个资料，要入门 Hartree Fock 方法，推荐看此书的第二章和第三章。
2. *Ideas of Quantum Chemistry*。从这书的附录中抄了两个公式。
3. 《手算量子化学》南开大学的张鋆写的一份量化编程资料。这是一份相当好的资料，值得仔细阅读。
4. 《闭壳层 H_n 分子电子能量编程教程》。参考了作者写的文档。