

HF Homework

Fernando Mata Rabadán

1 Introduction

1.1 Integral transformation from atomic to molecular orbitals

In order to evaluate the necessary matrix elements for the CIS and TD-HF algorithms, it is required to perform an integral transformation from atomic orbital basis to the molecular orbital basis:

$$(pq|rs) = \sum_{pqrs} \sum_{\mu\nu\lambda\sigma} c_{\mu p} c_{\nu q} c_{\lambda r} c_{\sigma s} (\mu\nu|\lambda\sigma)$$

This results in an algorithmic complexity of $\mathcal{O}(N^8)$ which will increase significantly as the basis set size expands. It is possible to reduce the computational cost considering that the coefficients $c_{\mu p}$, $c_{\nu q}$, $c_{\lambda r}$, and $c_{\sigma s}$ are independent with respect to the rest of the basis functions. Therefore, the previous expression can be rewritten as:

$$(pq|rs) = \sum_{\mu p} c_{\mu p} \left[\sum_{\nu q} c_{\nu q} \left[\sum_{\lambda r} c_{\lambda r} \left[\sum_{\sigma s} c_{\sigma s} (\mu\nu|\lambda\sigma) \right] \right] \right]$$

Then the integral transformation can be stepwise as:

$$\begin{aligned} (\mu\nu|\lambda s) &= \sum_{\sigma s} c_{\sigma s} (\mu\nu|\lambda\sigma) \implies (\mu\nu|rs) = \sum_{\lambda r} c_{\lambda r} (\mu\nu|\lambda s) \implies \\ \implies (\mu q|rs) &= \sum_{\nu q} c_{\nu q} (\mu\nu|rs) \implies (pq|rs) = \sum_{\mu p} c_{\mu p} (\mu q|rs) \end{aligned}$$

Which reduces the algorithmic complexity to $\mathcal{O}(4 \cdot N^5) \sim \mathcal{O}(N^5)$. The increase in speed of this approach is quite notable, especially when using larger basis sets.

1.2 Configuration interaction singles

In order to obtain the CIS energies, it is required to obtain and diagonalize the CIS matrix. The matrix elements of the CIS matrix are defined as:

$$A_{ai,jb} = \langle \Psi_j^b | \mathcal{H} - E_0 | \Psi_a^i \rangle = (\varepsilon_a - \varepsilon_i) \delta_{ij} \delta_{ab} + 2 \langle ib | aj \rangle - \langle ib | aj \rangle$$

Where i and j denote occupied orbitals and a b virtual ones, ε_i indicates the HF energy of the i orbital and $\langle ib | aj \rangle$ the electron repulsion integrals in the molecular orbital basis.

After building the CIS matrix, the energies and states can be obtained by diagonalization, being the eigenvalues the energies and the CIS states the eigenvectors.

1.3 Time dependent Hartree-Fock

The time dependent Hartree-Fock energies and states can be obtained by solving the equation:

$$\begin{pmatrix} A & B \\ -A & -B \end{pmatrix} \begin{pmatrix} X_m \\ Y_m \end{pmatrix} = \omega_m \begin{pmatrix} X_m \\ Y_m \end{pmatrix}$$

Where ω_w is the energy of each excited state and X_m and Y_m the excitation and de-excitation vectors. A and B are matrices defined as:

$$A_{ia,jb} = (\varepsilon_a - \varepsilon_i)\delta_{ij}\delta_{ab} + 2\langle ib|aj\rangle - \langle ib|aj\rangle$$

$$B_{ia,jb} = 2\langle ij|ab\rangle - \langle ib|ja\rangle$$

Where i and j denote occupied orbitals and a b virtual ones, ε_i indicates the HF energy of the i orbital and $\langle ib|aj\rangle$ the electron repulsion integrals in the molecular orbital basis.

The TDHF equation can be transformed into an eigenvalue problem as:

$$CZ_m = \omega_w^2 Z_m$$

Where the C matrix is defined as:

$$C = (A - B)^{1/2}(A - B)(A - B)^{1/2}$$

Where the square of the $(A - B)$ matrix was calculated via:

$$S^{1/2} = U S^{1/2} U^\dagger$$

Finally, by diagonalizing it is possible to obtain the TDHF energies as they will be the square root of the C matrix and the eigenvectors the TDHF states.

2 Code review

All code can be found in the GitHub repository: <https://github.com/matarabadanf/HF-Homework>.

2.1 Integral transformation

Following the second approach, results in four 5-fold nested **for** loops with the form:

```
do mu = 1, nBas
  do la = 1, nbas
    do nu = 1, nbas
      do s = 1, nbas
        do si = 1, nbas
          scr(mu, la, nu, s) = scr(mu, la, nu, s) + c(si, s) * eri_ao(mu, la,
            nu, si)
        end do
      end do
    end do
  end do
end do
```

Greatly reducing the computational cost.

2.2 CIS subroutine

To calculate the A matrix elements, it is necessary to determine all excitation pairs. To store all the possible excitations, an $nO \cdot nV \times 2$ dimension matrix is created and filled by iterating as:

```
do i = 1, no
  do j = no + 1, no + nv
    ex = ex + 1
    ex_vec(ex, 1) = i
    ex_vec(ex, 2) = j
  end do
end do
```

Knowing all pair of excitations, the A matrix elements are calculated with:

```
do ia = 1, ex
  do jb = 1, ex
    i = ex_vec(ia, 1)
    a = ex_vec(ia, 2)
    j = ex_vec(jb, 1)
    b = ex_vec(jb, 2)
    a_mat(ia, jb) = (e(a)-e(i)) * identity(i,j) * identity
      (a,b) + &
      2*(eri_mo(i,b,a,j)) - eri_mo(i,b,j,a)
  end do
end do
```

To obtain the CIS excitation energies and the character of each state, the A matrix is diagonalized with:

```
call diagonalize_matrix(ov, a_mat, cis_energies)
```

2.3 TDHF subroutine

Again the indices of the orbitals that are involved in all possible excitations are obtained, and the A matrix elements are calculated in the same way as in the previous section. The B matrix elements are calculated similarly with the formula definition.

It is required to compute the square root of the $(A - B)$ matrix in order to build the C matrix. This is done with:

```
a_p_b = a_mat + b_mat
a_m_b = a_mat - b_mat

call diagonalize_matrix(ov, a_m_b, a_m_b_eigen)

diag_a_m_b = 0.d0

do i = 1, ov
  diag_a_m_b(i,i) = a_m_b_eigen(i)
end do

a_m_b_r = matmul(a_m_b, matmul(sqrt(diag_a_m_b), transpose(a_m_b)))
```

Once the square root is obtained, the C matrix is constructed:

```
c_mat = matmul(a_m_b_r, matmul(a_p_b, a_m_b_r))
```

And through diagonalizing the C matrix it is possible to obtain the TDHF energies and states:

```
call diagonalize_matrix(ov, c_mat, TDHF_energies)

TDHF_energies = sqrt(TDHF_energies)
```

3 Results

With the previously written subroutines, the CIS and TDHF energies were calculated for the He , Ne , Be , and H_2O molecules using the cc-pVDZ and cc-pVTZ basis sets. The results are shown in Table 1. CIS energies were compared with the ones obtained using ORCA, providing quite similar results, validating the algorithm implementation.

Molecule / basis set	CIS	ORCA CIS	TDHF	NIST [1]
He / cc-pVDZ	51.947	51.947	51.577	20.615774
He / cc-pVTZ	31.807	32.819	31.608	20.615774
He / cc-pVDZ	28.162	29.090	28.018	20.615774
Ne / cc-pVDZ	49.001	49.008	48.871	16.619070
Ne / cc-pVTZ	33.210	36.950	33.174	16.619070
Be / cc-pVDZ	5.295	5.295	4.993	2.724963
Be / cc-pVTZ	5.143	5.148	4.867	2.724963
H_2O / cc-pVDZ	9.180	9.334	9.121	-

Table 1: 1st excitation energies in eV.

Even though the algorithms are correctly implemented, it can be seen that they present clear pitfalls when compared to the experimental results. While there is an improvement in the energy values when expanding the basis set, the error in energies is considerable, especially in Ne and He , with differences of 16 and 31 eV respectively.

The methodology used to compute excitation states is flawed due to the description of electron correlation. To recover electron correlation, it is necessary to resort to double excitations, but neither CIS nor TDHF consider them.

The CIS method only considers single excitations, which limits its ability to describe excited states accurately, especially for systems where electron correlation plays a significant role. This is evident from the discrepancies observed between the calculated CIS energies and the experimental values. Similarly, the TDHF method, although it includes both excitation and de-excitation vectors, still is not enough to properly describe electron correlation. The TDHF energies show some very slight improvements over CIS, but the differences with experimental values are still large. While this is an oversimplified explanation of these methods, for more detail we refer to the literature [2] or [3].

4 Conclusions

Throughout this homework, the CIS and TDHF methods were implemented in FORTRAN to calculate excitation energies for the He , Ne , Be , and H_2O molecules using the cc-pVDZ and cc-pVTZ basis sets. The results obtained were compared with experimental data and with the results obtained using ORCA, showing a general disagreement with experimental results and a great dependence in basis set size.

While the CIS and TDHF methods provide a useful starting point for studying excitation states, their limitations in describing electron correlation must be acknowledged. For more accurate results, methods that account for double excitations (or further) should be employed.

References

- (1) NIST Atomic Spectra Database, <https://dx.doi.org/10.18434/T4W30F>, November 2024.
- (2) Szabo, A.; Ostlund, N., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Books on Chemistry; Dover Publications: 1996.
- (3) Dreuw, A.; Head-Gordon, M. Single-reference ab initio methods for the calculation of excited states of large molecules, 2005.