

# *Electronic Structure Theory*

## *TSTC Session 8*



1. Born-Oppenheimer approx.- energy surfaces
2. Mean-field (Hartree-Fock) theory- orbitals
3. Pros and cons of HF- RHF, UHF
4. Beyond HF- why?
5. First, one usually does HF-how?
6. Basis sets and notations
7. MPn, MCSCF, CI, CC, DFT
8. **Gradients and Hessians**
9. Special topics: accuracy, metastable states

*Jack Simons*, *Henry Eyring Scientist and Professor*  
*Chemistry Department*  
*University of Utah*

The use of *analytical derivatives of the energy* with respect to atomic positions has made evaluation of *vibrational frequencies* and the mapping out of *reaction paths* much easier.

The first derivative with respect to a Cartesian coordinate ( $X_K$ ) of an atom is called the *gradient*

$$g_K = \partial E / \partial X_K.$$

These numbers form the gradient vector.

The second derivatives  $\partial^2 E / \partial X_K \partial X_L$  form the *Hessian matrix*  $H_{K,L}$

In the old days,  $g_K$  and  $H_{K,L}$  were evaluated by “finite difference” (evaluating the energy at slightly displaced  $X_K$ ).

Today, we have *analytical expressions* for  $g_K$  and  $H_{K,L}$ .

Two issues: How does one compute  $\mathbf{g}_K$  and  $H_{K,L}$  and what do you do with them?

Assume you have  $\mathbf{g}_K$  available at some starting geometry

$$\mathbf{X}^0 = \{X_1, X_2, \dots, X_{3N}\}.$$

One can attempt to move downhill toward a local-minimum by taking small displacements  $\delta\mathbf{X}_K$  proportional to, but in opposition to, the gradient  $\mathbf{g}_K$  along that direction

$$\delta\mathbf{X}_K = -a \mathbf{g}_K.$$

The energy  $E$  is then expected to change by

$$\delta E = -a \sum_K (\mathbf{g}_K)^2.$$

This is the most simple algorithm for “stepping” downhill toward a minimum.

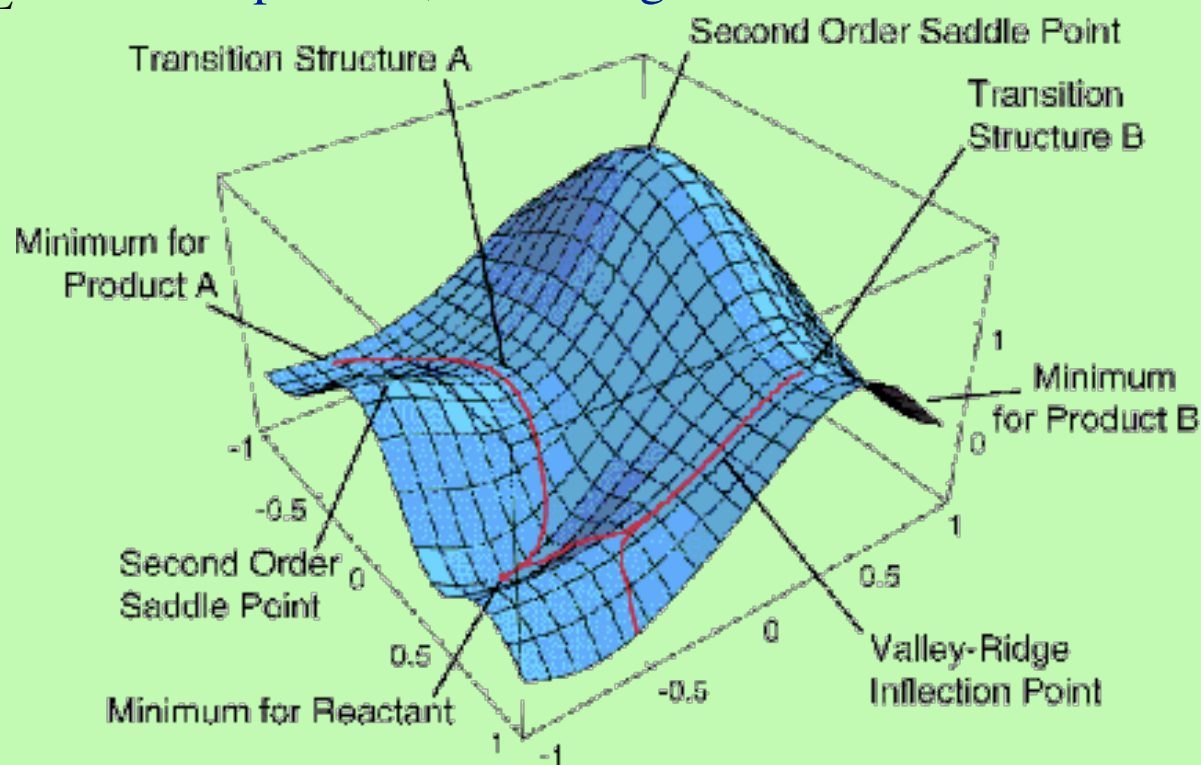
The parameter  $a$  can be used to keep the length of the step small.

A series of such “steps” from  $\mathbf{X}^0$  to  $\mathbf{X}^0 + \delta\mathbf{X}$  can often lead to a minimum (at which all  $3N$   $\mathbf{g}_K$  values vanish).

One problem with this approach is that, if one reaches a point where all  $3N$   $g_K$  vanish, one can not be certain it is a minimum; maybe it is a first-, second-, or higher-order saddle point.

**Minimum:** all  $3N$   $g_K$  vanish and  $3N-6$  eigenvalues of the  $H_{K,L}$  matrix are positive.

**First-order saddle (transition state TS):** all  $3N$   $g_K$  vanish and  $3N-7$  eigenvalues of the  $H_{K,L}$  matrix are positive; one is negative.



So, one is usually forced to form  $H_{K,L}$  and find its  $3N$  eigenvalues  $\lambda_a$  and eigenvectors  $V_k^a$

$$\sum_{L=1,3N} H_{K,L} V_L^a = \lambda_a V_k^a.$$

3 of the  $\lambda_a$  have to vanish and the 3 corresponding  $V_k^a$  describe translations of the molecule.

3 more (only 2 for linear molecules) of the  $\lambda_a$  have to vanish and the corresponding  $V_k^a$  describe rotations of the molecule.

The remaining  $3N-6$  (or  $3N-5$ )  $\lambda_a$  and  $V_k^a$  contain the information one needs to characterize the vibrations and reaction paths of the molecule.

If one has the **gradient vector** and **Hessian matrix** available at some geometry,

$$\delta E = \sum_K g_K \delta X_K + 1/2 \sum_{K,L} H_{K,L} \delta X_K \delta X_L$$

Because the Hessian is symmetric, its eigenvectors are **orthogonal**

$$\sum_K V_K^a V_K^b = \delta_{a,b}$$

and they form a **complete set**

$$\sum_a V_K^a V_L^a = \delta_{K,L}.$$

This allows one to express the atomic **Cartesian displacements**  $\delta X_K$  in terms of **displacements**  $\delta V^a$  along the “**eigenmodes**”

$$\delta X_K = \sum_L \delta_{K,L} \delta X_L = \sum_a V_K^a (\sum_L V_L^a \delta X_L) = \sum_a V_K^a \delta V^a.$$

Inserting

$$\delta X_K = \sum_a V_K^a \delta V^a.$$

into

$$\delta E = \sum_K g_K \delta X_K + 1/2 \sum_{K,L} H_{K,L} \delta X_K \delta X_L$$

gives

$$\delta E = \sum_a \{g_a \delta V^a + 1/2 \lambda_a (\delta V^a)^2\}$$

where

$$g_a = \sum_L V_L^a g_L$$

This way of writing  $\delta E$  allows us to consider independently **maximizing or minimizing** along **each** of the  $3N-6$  **eigenmodes**.

Setting the derivative of

$$\{g_a \delta V^a + 1/2 \lambda_a (\delta V^a)^2\}$$

with respect to the  $\delta V^a$  displacements equal to zero gives as a suggested “step”

$$\delta V^a = -g_a/\lambda_a$$

Inserting these displacements into

$$\delta E = \sum_a \{g_a \delta V^a + 1/2 \lambda_a (\delta V^a)^2\}$$

gives

$$\delta E = \sum_a \{-g_a^2/\lambda_a + 1/2 \lambda_a (-g_a/\lambda_a)^2\} = -1/2 \sum_a g_a^2/\lambda_a.$$

So the energy will go “downhill” along an eigenmode if that mode’s eigenvalue  $\lambda_a$  is positive; it will go uphill along modes with negative  $\lambda_a$  values.

Once you have a value for  $\delta V^a$ , you can compute the Cartesian displacements from

$$\delta X_K = \sum_a V_K^a \delta V^a$$



If one wants to find a minimum, one can

- a. Take a displacement  $\delta V^a = -g_a/\lambda_a$  along any mode whose  $\lambda_a$  is positive.
- b. Take a displacement that is **small and of opposite sign** than  $-g_a/\lambda_a$  for modes with negative  $\lambda_a$  values.

The energy will then decrease along all  $3N-6$  modes.

*What about finding transition states?*

## *What about finding transition states?*

If one is already at a geometry where one  $\lambda_a$  is negative and the  $3N-7$  other  $\lambda_a$  values are positive, one should

- Visualize the eigenvector  $\mathbf{V}_k^a$  belonging to the negative  $\lambda_a$  to make sure this displacement “makes sense” (i.e., looks reasonable for motion away from the desired transition state).
- If the mode having negative eigenvalue makes sense, one then takes

$$\delta V^a = -g_a/\lambda_a \text{ for all modes.}$$

This choice will cause

$$\delta E = \sum_a \left\{ -g_a^2/\lambda_a + 1/2 \lambda_a (-g_a/\lambda_a)^2 \right\} = -1/2 \sum_a g_a^2/\lambda_a$$

to go downhill along  $3N-7$  modes and uphill along the one mode having negative  $\lambda_a$ . Following a series of such steps may allow one to locate the TS at which all  $g_a$  vanish,  $3N-7$   $\lambda_a$  are positive and one  $\lambda_a$  is negative.

At a minimum or TS, one can evaluate **harmonic vibrational frequencies** using the Hessian. The gradient ( $g_L$  or  $g_a = \sum_L V_L^a g_L$  vanishes), so the local potential energy can be expressed in terms of the Hessian only.

The **classical dynamics Hamiltonian** for displacements  $\delta X_K$  is

$$H = \sum_{K,L} 1/2 H_{K,L} \delta X_K \delta X_L + 1/2 \sum_K m_K (d\delta X_K/dt)^2$$

Introducing the **mass-weighted Cartesian coordinates**

$$\delta MWX_K = (m_K)^{1/2} \delta X_K$$

allows the Hamiltonian to become

$$H = \sum_{K,L} 1/2 MWH_{K,L} \delta MWX_K \delta MWX_L + 1/2 \sum_K (d\delta MWX_K/dt)^2$$

where the **mass-weighted Hessian** is defined as

$$MWH_{K,L} = H_{K,L} (m_K m_L)^{-1/2}$$

Expressing the Cartesian displacements in terms of the eigenmode displacements

$$\delta X_K = \sum_a V_{K^a} \delta V^a$$

allows H to become

$$H = \sum_a \left\{ \frac{1}{2} \lambda_a (\delta V^a)^2 + \frac{1}{2} (d\delta V^a/dt)^2 \right\}.$$

This is the Hamiltonian for  $3N-6$  uncoupled harmonic oscillators having force constants  $\lambda_a$  and having unit masses for all coordinates. Thus, the harmonic vibrational frequencies are given by

$$\omega_a = (\lambda_a)^{1/2}$$

so the eigenvalues of the mass-weighted Hessian provide the harmonic vibrational frequencies.

At a TS, one of the  $\lambda_a$  will be negative.

It is worth pointing out that one can use mass-weighted coordinates to locate minima and transition states, but the same minima and transition states will be found whether one uses Cartesian or mass-weighted Cartesian coordinates because **whenever the Cartesian gradient vanishes**

$$g_L = 0$$

the **mass-weighted gradient will also vanish**

$$g_a = \sum_L V_L^a g_L = 0$$

To trace out a reaction path starting at a transition state, one first finds the Hessian eigenvector  $\{V_K^1\}$  belonging to the negative eigenvalue. One takes a very small step along this direction.

Next, one re-computes the Hessian and gradient (n.b., the gradient vanishes at the transition state, but not once begins to move along the reaction path) at the new geometry  $X_K + \delta X_K$  where one finds the eigenvalues and eigenvectors of the mass-weighted Hessian and uses the local quadratic approximation

$$\delta E = \sum_a \{g_a \delta V^a + 1/2 \lambda_a (\delta V^a)^2\}$$

to guide one downhill. Along the eigenmode corresponding to the negative eigenvalue  $\lambda_1$ , the gradient  $g_1$  will be non-zero while the components of the gradient along the other eigenmodes will be small (if one has taken a small initial step). One is attempting to move down a streambed whose direction of flow initially lies along  $V_K^1$  and perpendicular to which there are harmonic sidewalls  $1/2 \lambda_a (\delta V^a)^2$ .

One performs a series of displacements by

- a) moving (in small steps) downhill along the eigenmode that begins at  $V_K^1$  and that has a significant gradient component  $g_a$ ,
- b) while minimizing the energy (to remain in the streambed's bottom) along the  $3N-7$  other eigenmodes (by taking steps

$$\delta V^a = - g_a / \lambda_a \text{ that minimize each } \{g_a \delta V^a + 1/2 \lambda_a (\delta V^a)^2\}.$$

As one evolves along this reaction path, one reaches a point where  $\lambda_1$  changes sign from negative to positive. This signals that one is approaching a minimum. Continuing onward, one reaches a point where the gradient's component along the step displacement vanishes and along all other directions vanishes. This is the local minimum that connects to the transition state at which the reaction path started.

One needs to also begin at the transition state and follow the other branch of the reaction path to be able to connect reactants, transition state, and products.

When tracing out reaction paths, one uses the mass-weighted coordinates because dynamical theories (e.g. the reaction-path Hamiltonian theory) are formulated in terms of motions in mass-weighted coordinates. The minima and transition states one finds using mass-weighted coordinates will be the same as one finds using conventional Cartesian coordinates. However, the paths one traces out will differ depending on whether mass-weighting is or is not employed.



So, how does one evaluate the gradient and Hessian analytically?

For methods such as SCF, CI, and MCSCF that compute the energy E as

$$E = \langle \psi | H | \psi \rangle / \langle \psi | \psi \rangle,$$

one makes use of the chain rule to write

$$\begin{aligned} \partial E / \partial X_K = \sum_I \partial E / \partial C_I \partial C_I / \partial X_K + \sum_{i\mu} \partial E / \partial C_{i\mu} \partial C_{i\mu} / \partial X_K \\ + \langle \psi | \partial H / \partial X_K | \psi \rangle / \langle \psi | \psi \rangle. \end{aligned}$$

For MCSCF,  $\partial E / \partial C_I$  and  $\partial E / \partial C_{i\mu}$  are zero.

For SCF  $\partial E / \partial C_{i\mu}$  are zero and  $\partial E / \partial C_I$  does not exist.

For CI,  $\partial E / \partial C_I$  are zero, but  $\partial E / \partial C_{i\mu}$  are not.

So, for some of these methods, one needs to solve “response equations” for

$$\partial E / \partial C_{i\mu}.$$

What is  $\langle \psi | \partial H / \partial X_K \psi \rangle / \langle \psi | \psi \rangle$ ?

$$\langle \psi | H | \psi \rangle = \sum_{L,J} C_L C_J \langle | \phi_{L1} \phi_{L2} \phi_{L3} \dots \phi_{LN} | H | \phi_{J1} \phi_{J2} \phi_{J3} \dots \phi_{JN} \rangle$$

and each of the Hamiltonian matrix elements is given via Slater-Condon rules in terms of 1- and 2- electron integrals

$$\langle \phi_a | T_e + V_{e,n} + V_{n,n} | \phi_m \rangle \text{ and } \langle \phi_a(1) \phi_1(2) | 1/r_{1,2} | \phi_m(1) \phi_1(2) \rangle$$

The only places the nuclear positions  $X_K$  appear are

1. in the basis functions appearing in  $\phi_J = \sum_{\mu} \chi_{\mu} C_{J,\mu}$  and
2. in  $V_{e,n} = - \sum_a Z_a e^2 / |\mathbf{r} - \mathbf{R}_A|$

So,

$$\langle \psi | \partial H / \partial X_K \psi \rangle$$

will involve  $\langle \phi_a | \partial V_{e,n} / \partial X_K | \phi_m \rangle$

as well as derivatives  $\partial / \partial X_K$  of the  $\chi_{\mu}$  appearing in

$$\langle \phi_a | T_e + V_{e,n} + V_{n,n} | \phi_m \rangle \text{ and in } \langle \phi_a(1) \phi_1(2) | 1/r_{1,2} | \phi_m(1) \phi_1(2) \rangle$$

$$\partial/\partial \mathbf{X}_A V_{e,n} = - \sum_a Z_a (\mathbf{x}-\mathbf{X}_A) e^2/|\mathbf{r}-\mathbf{R}_A|^3$$

When put back into  $\langle \phi_a | \partial V_{e,n} / \partial \mathbf{X}_K | \phi_m \rangle$  and into the Slater-Condon formulas, these terms give the **Hellmann-Feynman contributions** to the gradient. These are not “difficult” integrals, but they are new ones that need to be added to the usual 1- electron integrals.

The  $\partial/\partial \mathbf{X}_K$  derivatives of the  $\chi_v$  appearing in

$$\langle \chi_v | -1/2 \nabla^2 | \chi_\mu \rangle + \sum_a \langle \chi_v | -Z_a / r_a | \chi_\mu \rangle$$
 and in

$$\langle \chi_v(\mathbf{r}) \chi_\eta(\mathbf{r}') | (1/|\mathbf{r}-\mathbf{r}'|) | \chi_\mu(\mathbf{r}) \chi_\gamma(\mathbf{r}') \rangle$$

present major new difficulties because they involve new integrals

$$\langle \partial/\partial \mathbf{X}_K \chi_v | -1/2 \nabla^2 | \chi_\mu \rangle + \sum_a \langle \partial/\partial \mathbf{X}_K \chi_v | -Z_a / r_a | \chi_\mu \rangle$$

$$\langle \partial/\partial \mathbf{X}_K \chi_v(\mathbf{r}) \chi_\eta(\mathbf{r}') | (1/|\mathbf{r}-\mathbf{r}'|) | \chi_\mu(\mathbf{r}) \chi_\gamma(\mathbf{r}') \rangle$$

When Cartesian Gaussians

$$\chi_{a,b,c}(\mathbf{r},\theta,\phi) = N'_{a,b,c,\alpha} x^a y^b z^c \exp(-\alpha r^2)$$

are used, the derivatives  $\partial/\partial X_K \chi_v(\mathbf{r})$  can be done because  $X_K$  appears in

$$(x-X_K)^a \text{ and in } r^2 = (x-X_K)^2 + (y-Y_K)^2 + (z-Z_K)^2 .$$

These derivatives give functions of **one lower**

$$\text{(from } \partial/\partial X_K (x-X_K)^a \text{)}$$

and **one higher**

$$\text{(from } \partial/\partial X_K \exp(-\alpha r^2)\text{)}$$

**angular momentum** value.

So, the AO integral list must be extended to higher L-values.

More troublesome are

$$\langle \partial/\partial X_K \chi_v(\mathbf{r}) \chi_\eta(\mathbf{r}') | (1/|\mathbf{r}-\mathbf{r}'|) | \chi_\mu(\mathbf{r}) \chi_\gamma(\mathbf{r}') \rangle$$

because there are now **4 times** ( the original plus  $\partial/\partial X_K, \partial/\partial Y_K, \partial/\partial Z_K$ ) the **number of 2-electron integrals**.

When **plane wave basis** functions are used, the derivatives

$$\partial/\partial X_K \chi_v(\mathbf{r}) = 0$$

vanish (and thus don't have to be dealt with) because the basis functions do not "sit" on any particular nuclear center. This is a substantial benefit to using plane waves.

The good news is that the **Hellmann-Feynman** and integral derivative terms **can be evaluated** and thus the gradients can be computed as

$$\begin{aligned} \partial E / \partial X_K &= \sum_I \partial E / \partial C_I \partial C_I / \partial X_K + \sum_{i\mu} \partial E / \partial C_{i\mu} \partial C_{i\mu} / \partial X_K \\ &+ \langle \psi | \partial H / \partial X_K | \psi \rangle / \langle \psi | \psi \rangle = \langle \psi | \partial H / \partial X_K | \psi \rangle / \langle \psi | \psi \rangle \end{aligned}$$

for **SCF** or **MCSCF** wavefunctions.

*What about CI, MPn, or CC wave functions? What is different?*

$$\frac{\partial E}{\partial X_K} = \sum_I \frac{\partial E}{\partial C_I} \frac{\partial C_I}{\partial X_K} + \sum_{i\mu} \frac{\partial E}{\partial C_{i\mu}} \frac{\partial C_{i\mu}}{\partial X_K} + \frac{\langle \psi | \partial H / \partial X_K | \psi \rangle}{\langle \psi | \psi \rangle}.$$

- For **CI**, the  $\frac{\partial E}{\partial C_I}$  term still vanishes and the

$$\frac{\langle \psi | \partial H / \partial X_K | \psi \rangle}{\langle \psi | \psi \rangle}$$

term is handled as in MCSCF, but the  $\frac{\partial E}{\partial C_{i\mu}}$  terms do not vanish

- For **MPn**, one does not have  $C_I$  parameters; E is given in terms of orbital energies  $\epsilon_j$  and 2-electron integrals over the  $\phi_j$ .
- For **CC**, one has  $t_{i,j}^{m,n}$  amplitudes as parameters and E is given in terms of them and integrals over the  $\phi_j$ .

So, in **CI**, **MPn**, and **CC** one needs to have expressions for

$$\frac{\partial C_{i\mu}}{\partial X_K} \text{ and for } \frac{\partial t_{i,j}^{m,n}}{\partial X_K}.$$

These are called *response equations*.

The response equations for  $\partial C_{i,\mu} / \partial X_K$  are obtained by taking the  $\partial / \partial X_K$  derivative of the Fock equations that determined the  $C_{i,\mu}$

$$\partial / \partial X_K \sum_{\mu} \langle \chi_{\nu} | h_e | \chi_{\mu} \rangle C_{J,\mu} = \partial / \partial X_K \varepsilon_J \sum_{\mu} \langle \chi_{\nu} | \chi_{\mu} \rangle C_{J,\mu}$$

This gives

$$\begin{aligned} & \sum_{\mu} [\langle \chi_{\nu} | h_e | \chi_{\mu} \rangle - \varepsilon_J \langle \chi_{\nu} | \chi_{\mu} \rangle] \{ \partial / \partial X_K C_{J,\mu} \} = \\ & - \sum_{\mu} \partial / \partial X_K [\langle \chi_{\nu} | h_e | \chi_{\mu} \rangle - \varepsilon_J \sum_{\mu} \langle \chi_{\nu} | \chi_{\mu} \rangle] C_{J,\mu} \end{aligned}$$

Because all the machinery to evaluate the terms in

$$\partial / \partial X_K [\langle \chi_{\nu} | h_e | \chi_{\mu} \rangle - \varepsilon_J \sum_{\mu} \langle \chi_{\nu} | \chi_{\mu} \rangle]$$

exists as does the matrix

$$\langle \chi_{\nu} | h_e | \chi_{\mu} \rangle - \varepsilon_J \langle \chi_{\nu} | \chi_{\mu} \rangle,$$

one can solve for  $\partial / \partial X_K C_{J,\mu}$



A similar, but more complicated, strategy can be used to derive equations for the  $\partial t_{ij}^{m,n} / \partial X_K$  that are needed to achieve gradients in CC theory.

The bottom line is that for MPn, CI, and CC, one can obtain analytical expressions for  $g_K = \partial E / \partial X_K$ .

To derive analytical expressions for the Hessian  $\partial^2 E / \partial X_K \partial X_L$  is, of course, more difficult. It has been done for HF and MCSCF and CI and may exist (?) for CC theory. As you may expect it involves second derivatives of 2-electron integrals and thus is much more “expensive”.

There are **other kinds of responses** that one can seek to treat analytically. For example, what if one added to the Hamiltonian an electric field term such as  $\sum_{k=1,N} \mathbf{e} \mathbf{r}_k \bullet \mathbf{E} + \sum_{a=1,M} e Z_a \mathbf{R}_a \bullet \mathbf{E}$  rather than displacing a nucleus?

So, H is now  $H + \sum_{k=1,N} \mathbf{e} \mathbf{r}_k \bullet \mathbf{E} + \sum_{a=1,M} e Z_a \mathbf{R}_a \bullet \mathbf{E}$ .

The wavefunction  $\psi(\mathbf{E})$  and energy  $E(\mathbf{E})$  will now depend on the electric field  $\mathbf{E}$ .

$$dE/d\mathbf{E} = \sum_I \partial E / \partial C_I \partial C_I / \partial \mathbf{E} + \sum_{i\mu} \partial E / \partial C_{i\mu} \partial C_{i\mu} / \partial \mathbf{E} + \langle \psi | \partial H / \partial \mathbf{E} | \psi \rangle / \langle \psi | \psi \rangle.$$

$$\text{Here, } \langle \psi | \partial H / \partial \mathbf{E} | \psi \rangle / \langle \psi | \psi \rangle = \langle \psi | \sum_{k=1,N} \mathbf{e} \mathbf{r}_k + \sum_{a=1,M} e Z_a \mathbf{R}_a | \psi \rangle,$$

is the dipole moment expectation value. This is the **final answer for HF and MCSCF**, but **not for MPn, CI, CC**.

For these cases, we also need  $\partial C_I / \partial \mathbf{E}$  and  $\partial C_{i\mu} / \partial \mathbf{E}$  response contributions. So, the expectation value of the dipole moment operator is not always the correct dipole moment!