Breakdown of the Hellmann-Feynman theorem: Degeneracy is the key

G. P. Zhang
Department of Physics, Indiana State University, Terre Haute, Indiana 47809

Thomas F. George
Office of the Chancellor/Departments of Chemistry and Physics & Astronomy, University of Wisconsin-Stevens Point, Stevens Point, Wisconsin 54481-3897

(Received 17 May 2002; published 29 July 2002)

The Hellmann-Feynman theorem is a powerful and popular method to efficiently calculate forces in a variety of dynamical processes, but its validity has rarely been addressed. Here a surprising failure of this theorem is reported. The forces calculated by the theorem can be more than fifty times smaller than the forces calculated by the finite differential method. Numerical evidence shows that the energy-level degeneracy is the main reason. An analytical proof reveals that although eigenvalues do not depend on a linear combination of degenerate wave functions, forces do sensitively depend on it, which leads to ill-defined forces. A scheme is proposed to overcome this difficulty.

DOI: 10.1103/PhysRevB.66.033110 PACS number(s): 71.15.−m

The Hellmann-Feynman theorem[1] is powerful and has been widely used in many fields such as dynamical processes,[2] molecular dynamics,[3] chemical reactions, and surface reconstructions.[4–6] The beauty of this theorem is that in the adiabatic limit, it requires the diagonalization of Hamiltonian matrices only once and enables one to calculate all forces without recalculating wave functions. This saves a tremendous amount of computational power, in particular in those heavy-duty calculations such as *ab initio* calculations.[2] This appealing feature has motivated extensive investigations for a long time,[5] but the validity of the Hellmann-Feynman theorem has been rarely addressed in the literature.

In this paper, we report a surprising failure of the Hellmann-Feynman theorem. Forces computed by the theorem are inconsistent with those computed by the finite differential method, and even the signs of forces are totally different. We find that the energy-level degeneracy is the main reason. In the presence of the degeneracy, the force given by the Hellmann-Feynman theorem is incorrect. Numerical examples for C_{60} explicitly show that forces calculated by the Hellmann-Feynman theorem can be fifty times smaller than forces calculated by the finite differential method, though the finite differential method itself also gives inconsistent results. An analytical calculation shows that although any combination of degenerate eigenstates are still eigenstates of the system, forces become ill defined. Therefore, it is dangerous to use those forces to do simulations. A scheme is proposed to overcome this difficulty.

We consider a physical system which is described by the Hamiltonian matrix H. Such a system can be a solid,[2] liquid,[3] or molecule system.[1] Within the adiabatic approximation, the electron follows the motion of atoms. By diagonalizing the Hamiltonian matrix, we obtain both the eigenvectors $\{|\Psi_n\rangle\}$ and eigenvalues $\{E_n\}$, where $H|\Psi_n\rangle = E_n|\Psi_n\rangle$, or $E_n = \langle\Psi_n|H|\Psi_n\rangle$ and $\langle\Psi_n|\Psi_m\rangle = \delta_{nm}$. The force $\vec{F}_{n,\beta}$ corresponding to level n can be computed by taking the derivative of the energy E_n with respect to the position $\{\vec{r}_\beta\}$ of the atoms along the direction β,

$$\vec{F}_{n,\beta} = -\frac{\partial E_n}{\partial \vec{r}_\beta} = -\frac{\partial}{\partial \vec{r}_\beta}\langle\Psi_n|H|\Psi_n\rangle = -\left(\frac{\partial}{\partial \vec{r}_\beta}\langle\Psi_n|\right)H|\Psi_n\rangle - \left\langle\Psi_n\left|\frac{\partial H}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle - \left\langle\Psi_n\left|H\left(\frac{\partial}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle\right)$$
$$= -\left(\frac{\partial}{\partial \vec{r}_\beta}\langle\Psi_n|\right)E_n|\Psi_n\rangle - \left\langle\Psi_n\left|\frac{\partial H}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle - \left\langle\Psi_n\left|E_n\left(\frac{\partial}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle\right) = -\left\langle\Psi_n\left|\frac{\partial H}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle - E_n\frac{\partial}{\partial \vec{r}_\beta}(\langle\Psi_n|\Psi_n\rangle). \quad (1)$$

Using the normalization condition $\langle\Psi_n|\Psi_n\rangle = 1$, we can simplify Eq. (1) as

$$\vec{F}_{n,\beta} = -\left\langle\Psi_n\left|\frac{\partial H}{\partial \vec{r}_\beta}\right|\Psi_n\right\rangle. \quad (2)$$

This is the Hellmann-Feynman theorem. Throughout the derivation, except for the adiabatic approximation, we have not invoked any assumption, so that the scheme is fully generic. The important feature of this theorem is that we can calculate forces for different directions after we diagonalize the Hamiltonian matrix only once. In real calculations, this is very attractive.

Here, we take C_{60} as an example.[7–9] Note that our conclusion is independent of such a selection.[10] C_{60} has the high-

TABLE I. Comparison of forces calculated by the finite differential method and the Hellmann-Feynman theorem. The second and third columns are forces by the finite differential method and the Hellmann-Feynman theorem, respectively. In the fifth and sixth columns are their average forces computed by the proposed scheme, where a good agreement can be seen clearly.

Level index	F_{finite} (eV/Å)	F_{HF} (eV/Å)	$F_{\text{finite}}/F_{\text{HF}}$	$\overline{F}_{\text{finite}}$ (eV/Å)	$\overline{F}_{\text{HF}}$ (eV/Å)	Energy (eV)
1	0.0403870	0.0403871	1.00	0.0403871	0.0403870	-6.6675130
2	-0.1244360	-0.1244359	1.00	-0.1244359	-0.1244360	-6.6573438
3	-0.0808800	-0.0808799	1.00	-0.0808799	-0.0808800	-6.1291991
4	-0.1408974	0.0528522	-2.67			-6.1095001
5	-0.1409309	-0.3346804	0.42	-0.1409141	-0.1409141	-6.1095001
6	-0.0798105	-0.3211777	0.25			-5.1194132
7	-0.0796447	0.1617228	-0.49	-0.0797275	-0.0797276	-5.1194132
8	-0.0457316	-0.0457315	1.00	-0.0457315	-0.0457316	-5.1099250
9	-0.1971248	-0.2476372	0.80			-5.0878820
10	-0.1973077	-0.1467953	1.34	-0.1972162	-0.1972163	-5.0878820
11	0.0656571	0.0656579	1.00	0.0656579	0.0656571	-4.0030869
12	-0.3895717	-0.7615671	0.51			-3.9876889
13	-0.3897314	-0.0177367	21.97	-0.3896519	-0.3896516	-3.9876889
14	0.2610102	0.0513534	5.08			3.5868688
15	0.1162016	0.3542612	0.33			3.5868688
16	0.1162017	0.3301816	0.35			3.5868688
17	0.1162016	-0.0022564	-51.50			3.5868688
18	0.2610740	0.1371496	1.90	0.1741379	0.1741378	3.5868688

est point symmetry of I_h with 60 carbon atoms situated at the vertices of the ball. There are a total of 60 π electrons on the ball.[11] The Hamiltonian for the whole system can be written as[12]

$$H=-\sum_{i,j,\sigma} t_{ij}(c^{\dagger}_{i,\sigma}c_{j,\sigma}+\text{H.c.})+\frac{K}{2}\sum_{i,j}(|\vec{r}_i-\vec{r}_j|-d_0)^2. \tag{3}$$

Here, $c^{\dagger}_{i\sigma}$ creates a π electron at site i with spin projection σ; $t_{ij}=t_0-\alpha(|\vec{r}_i-\vec{r}_j|-d_0)$ is the hopping integral between nearest-neighbor atoms at $\vec{r}_i$ and $\vec{r}_j$, where t_0 is the average hopping constant, α is the electron-lattice coupling constant, and $d_0=1.54$ Å; and K is the spring constant. The second term in Eq. (3) is the lattice elastic energy. By fitting the optical energy gap and two different bond lengths, we have determined the above parameters as $t_0=1.8$ eV, $\alpha=3.5$ eV/Å and $K=30.0$ eV/Å^2.[13–16]

We use both the regular finite differential method and the Hellmann-Feynman theorem to calculate the force. In the finite differential method, we choose a small deviation $\delta=10^{-4}$ Å to compute the force,

$$\vec{F}_{n,\beta,\text{finite}}=-\frac{E_n(\vec{r}_\beta+\vec{\delta})-E_n(\vec{r}_\beta-\vec{\delta})}{2\delta}. \tag{4}$$

In Table I, we compare the forces calculated by these two methods, where F_{finite} as calculated by the finite differential method are listed in the second column, while F_{HF} by the Hellmann-Feynman theorem are in the third column. It is very impressive that for energy level 1, these two forces agree up to the seventh digit with an error of 10^{-7} eV/Å. For level 2, we also see extremely good agreement, and the signs of the forces are the same. The same is true for level 3. Such a high accuracy is numerical proof of the validity of the Hellmann-Feynman theorem.

However, when we compare the results for level 4, there is a surprise. We notice that F_{finite} and F_{HF} do *not* agree with each other, where F_{finite} is almost three times larger than F_{HF}, and even their signs are different. If we were to use this force to do numerical simulations, we would make a serious mistake. For level 5, we see the inconsistency again, though this time the signs of forces are same. For levels 6 and 7, the forces by these two schemes are also different. But, for level 8, we see the good agreement again. Such agreements and disagreements are puzzling. More puzzling is that such a trend persists for other levels from 9 to 18. The biggest discrepancy is for level 17, where one sees that the force calculated by the Hellmann-Feynman theorem is more than 50 times smaller than the force calculated by the finite differential method. The ratio between F_{finite} and F_{HF} is listed in the fourth column of Table I. To resolve such a puzzle, we have carefully checked all numerical steps and can exclude any numerical errors. On the other hand, from the above derivation of the Hellmann-Feynman theorem, we know our scheme is purely generic, and there is no hidden assumption at all. This raises a serious question as to why the Hellmann-Feynman theorem fails.

After an extensive and careful examination, we find a systematic pattern in those energy levels whose forces are not consistent. In the last column of Table I, we list their eigenvalues. We found that those energy levels with inconsistent forces have at least twofold degeneracies. For instance, since levels 4 and 5 are degenerate, none of their

forces calculated by the Hellmann-Feynman theorem are consistent with those by the finite differential method. The same also holds true for levels 6 and 7. For level 8, there is no degeneracy, and we find the two forces are consistent, but for levels 9 and 10, we have the degeneracy and lose the good agreement again. Levels 14 to 18 have fivefold degeneracies, and F_{finite} and F_{HF} are not consistent with respect to each other.

Those numerical results suggest a possible breakdown of the Hellmann-Feynman theorem due to the level degeneracy. Next, we prove this analytically. Assume two degenerate levels of energy E with wave functions as

$$|\Psi_1\rangle = a|\phi_1\rangle + b|\phi_2\rangle, \tag{5}$$

$$|\Psi_2\rangle = c|\phi_1\rangle + d|\phi_2\rangle, \tag{6}$$

where $|\phi_1\rangle$ and $|\phi_2\rangle$ are two basis functions, and a, b, c, and d are real coefficients, to be consistent with our numerical calculations. From the orthornormalization, we have $a^2+b^2=1$, $c^2+d^2=1$, $ac+bd=0$. Therefore, the force $\vec{F}_{1(2)}$ for $|\Psi_{1(2)}\rangle$ can be calculated from the Hellmann-Feynman theorem,

$$\vec{F}_1 = -\left\langle \Psi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \Psi_1 \right\rangle = -a^2 \left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle - b^2 \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle - ab \left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle - ab \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle, \tag{7}$$

$$\vec{F}_2 = -\left\langle \Psi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \Psi_2 \right\rangle = -c^2 \left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle - d^2 \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle - cd \left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle - cd \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle. \tag{8}$$

Since the states $|\Psi_1\rangle$ and $|\Psi_2\rangle$ are degenerate, any linear combination of these two states are still the eigenstates of Hamiltonian H. In other words, we can form a new pair of wave functions like $|\Psi_1'\rangle = a'|\Psi_1\rangle + b'|\Psi_2\rangle$ and $|\Psi_2'\rangle = c'|\Psi_1\rangle + d'|\Psi_2\rangle$. Such a combination has no effect on the eigenvalue, but has a strong effect on the force. From Eqs. (7) and (8), one sees that the force sensitively depends on those coefficients. Therefore, for degenerate levels, the forces are ill defined. This explains why the forces calculated by different schemes may be different. In fact, even within the same finite differential scheme, the forces are also different (see, for instance, levels 4 and 5). For nondegenerate levels, since there is no such combination, the force is well defined, and results from the finite differential scheme and the Hellmann-Feynman scheme are consistent.

The next question is how to solve this problem. A simple way is to sum over all the forces which belong to the same degenerate levels. In the above example, we have

$$\begin{aligned} -(\vec{F}_1+\vec{F}_2) = {} & (a^2+c^2)\left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle + (b^2+d^2) \\ & \times \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle + (ab+cd)\left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle \\ & + (ab+cd)\left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle, \end{aligned} \tag{9}$$

which can be simplified to

$$-(\vec{F}_1+\vec{F}_2) = \left\langle \phi_1 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_1 \right\rangle + \left\langle \phi_2 \left| \frac{\partial H}{\partial \vec{r}} \right| \phi_2 \right\rangle, \tag{10}$$

where there is no more free parameter. This method works very well in reality. In particular, in the above numerical example, we find that for levels 4 and 5, the average force is 0.140 914 15 eV/Å, which is fully consistent with the average force from the Hellmann-Feynman theorem. In fact, using this method, all the forces in Table I computed by the finite differential method and the Hellmann-Feynman theorem are consistent with each other (compare columns 5 and 6). Therefore, we have provided a viable method to overcome the above difficulty.

In conclusion, we have found that the Hellmann-Feynman theorem breaks down in presence of the energy-level degeneracies. We have presented a numerical example showing that the failure is directly connected with the level degeneracy. The forces calculated by the Hellmann-Feynman theorem can be more than 50 times smaller than the forces calculated by the finite differential method, though the forces given by the finite differential method themselves are not consistent. Our analytic calculation proves that in the presence of energy-level degeneracies, any linear combination of the eigenstates are still eigenstates of the original system and have no effect on the eigenenergy, but they do affect the forces since the forces sensitively depend on such a combination, which leads to them being ill defined. A scheme to overcome this difficulty has been outlined, which works extremely well in reality. Our results have substantial impact on the present state-of-art numerical simulations, in particular, as the Hellmann-Feynman theorem has been so popularly used. We caution that those numerical forces may be erroneous.

[1] H. Hellmann, *Einfuhrung in die Quantenchemie* (Deuieke, Leipzig, 1937); R. P. Feynman, Phys. Rev. **56**, 340 (1939).
[2] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).
[3] M. P. Allen and D. J. Tildesley, *Computer Simulations of Liquids* (Clarendon, Oxford, 1987).
[4] P. E. Blöchl, Phys. Rev. B **50**, 17 953 (1994); M. C. Payne, M. P. Teter, D. C. Allan, T. A. Arias, and J. D. Joannopoulos, Rev. Mod. Phys. **64**, 1045 (1992).
[5] M. Weinert and J. W. Davenport, Phys. Rev. B **45**, 13 709 (1992); R. M. Wentzcovitch, J. L. Martins, and P. B. Allen, *ibid.* **45**, 11 372 (1992); J. L. Corkill and K.-M. Ho, *ibid.* **54**, 5340 (1996); F. M. Fernandez, Chem. Phys. Lett. **233**, 651 (1995).
[6] J. C. Tully, J. Chem. Phys. **93**, 1061 (1990); P. Pulay, Mol. Phys. **17**, 197 (1969).
[7] A. F. Hebard *et al.*, Nature (London) **350**, 600 (1991).
[8] K. Holczer *et al.*, Science **252**, 1154 (1991).
[9] S. Chakravarty, M. P. Gelfand, and S. Kivelson, Science **254**, 970 (1991).
[10] In some materials with lower symmetry geometries, the state degeneracy may be lower than those in C_{60}. But as far as one has level degeneracy, the force is not well defined. Note that such degeneracy can be accidental or regular. In each case, our formalism suggests a generic way to overcome such a difficulty while completely avoiding those unphysical partial occupations (see Ref. 5).
[11] Our conclusion is independent of the system that we are using. In our analytic example below, we only have a twofold degeneracy which already reveals the same conclusion. The Jahn-Teller effect may lift the state degeneracies, but not always. For instance, in linear molecules, according to the Jahn-Teller theorem, it cannot lower the symmetry.
[12] G. P. Zhang, R. T. Fu, X. Sun, D. L. Lin, and T. F. George, Phys. Rev. B **50**, 11 976 (1994).
[13] G. P. Zhang, X. Sun, T. F. George, and L. N. Pandey, J. Chem. Phys. **106**, 6398 (1997).
[14] G. P. Zhang, Y. S. Ma, X. Sun, K. H. Lee, and T. Y. Park, Phys. Rev. B **52**, 6081 (1995).
[15] X. Sun, G. P. Zhang, Y. S. Ma, K. H. Lee, T. Y. Park, T. F. George, and L. Pandey, Phys. Rev. B **53**, 15 481 (1996).
[16] R. T. Fu, G. P. Zhang, X. Sun, and T. F. George, J. Cluster Sci. **276**, 155 (1999).