Generalization of the Hellmann–Feynman theorem

J.G. Esteve [a,b,*], Fernando Falceto [a,b], C. García Canal [c]

[a] *Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain*
[b] *Instituto de Biocomputación y Física de Sistemas complejos (BIFI), Universidad de Zaragoza, 50009 Zaragoza, Spain*
[c] *Laboratorio de Física Teórica, Departamento de Física, Facultad de Ciencias Exactas, Universidad Nacional de La Plata and IFLP-CONICET, Argentina*

ARTICLE INFO

ABSTRACT

The well-known Hellmann–Feynman theorem of quantum mechanics connected with the derivative of the eigenvalues with respect to a parameter upon which the Hamiltonian depends, is generalized to include cases in which the domain of definition of the Hamiltonian of the system also depends on that parameter.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

The Hellmann–Feynman theorem (from now on the HF theorem) is a useful tool in solid state, atomic and molecular physics. One of its consequences is that in quantum mechanics there is a single way of defining a generalized force on eigenstates of the Hamiltonian, associated with the variation of some of its parameters.

Classically, given a Hamiltonian $U(\lambda)$ depending on a parameter λ (which can be a generalized coordinate), one can define a generalized force $F_\lambda = -\partial_\lambda U$ which is associated with that parameter in the sense that $F_\lambda \, d\lambda$ is the work done in changing the parameter by $d\lambda$. However, in quantum mechanics there are, in principle, two possible ways to implement that definition: given the quantum Hamiltonian $H(\lambda)$ (with eigenvalues $E_n(\lambda)$ and normalized eigenvectors $\Psi_n(\lambda)$), we can define the generalized force acting on the state $\Psi_n(\lambda)$ as $F_\lambda = -\partial_\lambda E_n(\lambda)$ or as the expectation value of $-\partial_\lambda H(\lambda)$ on the state $\Psi_n(\lambda)$, which gives $F_\lambda = \langle \Psi_n | -\partial_\lambda H(\lambda) \Psi_n \rangle$. The HF theorem ensures that both definitions are equivalent, i.e.

$$\frac{\partial E_n(\lambda)}{\partial \lambda} = \left\langle \Psi_n(\lambda) \left| \frac{\partial H(\lambda)}{\partial \lambda} \Psi_n(\lambda) \right. \right\rangle, \qquad (1)$$

where, obviously, the differentiability of E_n, H and Ψ_n with respect to λ is assumed. This equation is known as the differential form of the HF theorem and from it we can obtain the integrated version,

$$E_n(\lambda_1) - E_n(\lambda_2) = \frac{\langle \Psi_n(\lambda_2) | (H(\lambda_1) - H(\lambda_2)) \Psi_n(\lambda_1) \rangle}{\langle \Psi_n(\lambda_2) | \Psi_n(\lambda_1) \rangle}, \qquad (2)$$

and the off diagonal form,

$$\left(E_m(\lambda) - E_n(\lambda) \right) \left\langle \Psi_n(\lambda) \left| \frac{\partial}{\partial \lambda} \Psi_m(\lambda) \right. \right\rangle = \left\langle \Psi_n(\lambda) \left| \frac{\partial H}{\partial \lambda} \Psi_m(\lambda) \right. \right\rangle. \qquad (3)$$

Eqs. (1)–(3), and others derived from them have been used in many areas of physics and especially in solid state and molecular physics, with the pioneering work of Feynman [1] proving the so-called electrostatic theorem. Taking λ to be the coordinate X_i of the position of the nucleus i and assuming that there are no external fields, the HF theorem states that the force on nucleus i due to the other nuclei and the electrons, in some particular configuration, is exactly what could be computed in classical electrostatics from the location of the other nuclei and the electronic charge density. Furthermore, the use of the HF theorem has been extended to variational states that are not the true eigenstates of the Hamiltonian or even to Gamow states [2–4].

However, there are also systems for which the HF theorem fails for the eigenstates of the Hamiltonian, even if they fulfill the general conditions established for the validity of the theorem (differentiability with respect to the parameter). As we shall see later, these situations are related to the fact that in quantum systems the Hamiltonian can depend on a parameter, not only through an explicit dependence in the operator but also because the domain of definition of the Hamiltonian depends on that parameter. This happens, for instance, when the Hamiltonian contains interactions that are gauge invariant and then, the functional form of H and the boundary conditions for the functions of its domain both depend on the particular gauge we chose. In this case, as we shall see, the standard derivation of the HF theorem is not valid and a generalization is needed to cover those systems. This is precisely the main goal of this contribution.
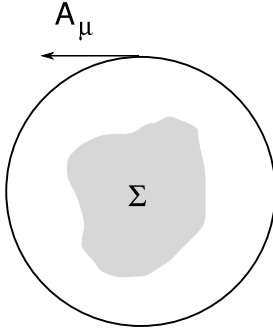
Fig. 1. The planar rotor under the influence of a magnetic field confined in the region Σ.

The Letter is organized as follows: In the next section we present a simple example of a Hamiltonian with magnetic interactions for which the HF theorem fails, in Section 3 we derive the generalization of the HF theorem, the application to some Hamiltonians, including the one mentioned above, is contained in Section 4, and finally we sketch the conclusions.

2. Particle interacting with a confined magnetic field

In this section we present an example of a system for which the standard form of the HF theorem (1) fails. The system describes a nonrelativistic charged particle moving on a circumference S^1 and interacting with a magnetic field which is confined in some region Σ on the interior of S^1 with magnetic flux equal to $2\pi\epsilon$ (see Fig. 1). This is a model for the superconducting Cooper pairs moving on a ring that is transversed by a magnetic flux $\Phi = 2\pi\epsilon$ [5].

Using suitable units, and in a gauge where the electromagnetic field A_μ is tangent to the circumference, with constant norm ϵ and directed in the anticlockwise direction (consequently the flux through Σ is $\Phi = \int_\Sigma \vec{B}\, d\vec{S} = \int_{S^1} \vec{A}\, d\vec{l} = 2\pi\epsilon$), the Hamiltonian can be written

$$H = -\frac{1}{2}\left(\frac{d}{d\theta} - i\epsilon\right)^2, \tag{4}$$

acting on periodic functions. The domain of definition of the Hamiltonian D_0 may be taken as the set of twice differentiable functions $f(\theta)$ (with $\theta \in [0, 2\pi]$), such that $f(0) = f(2\pi)$ and $\partial_\theta f(0) = \partial_\theta f(2\pi)$. In D_0, H is essentially selfadjoint and has the eigenvalues and normalized eigenvectors:

$$E_n(\epsilon) = \frac{1}{2}(n - \epsilon)^2, \tag{5}$$

$$\Psi_n(\theta) = \frac{1}{\sqrt{2\pi}}\exp(in\theta). \tag{6}$$

Note that the eigenvectors of the Hamiltonian do not depend on ϵ. In this gauge, $\partial_\epsilon E_n(\epsilon) = \epsilon - n$, $\partial_\epsilon H(\epsilon) = i\partial_\theta + \epsilon$ and the HF equation (1) is exactly satisfied

$$\frac{\partial E_n(\epsilon)}{\partial \epsilon} = \epsilon - n = \left\langle \Psi_n \left| \frac{\partial H(\epsilon)}{\partial \epsilon} \Psi_n \right. \right\rangle. \tag{7}$$

However, we could also work in another gauge obtained by the unitary transformation

$$H_g = e^{-i\epsilon\theta} H e^{i\epsilon\theta} = -\frac{1}{2}\frac{d^2}{d\theta^2}, \tag{8}$$

$$D_\epsilon = e^{-i\epsilon\theta} D_0, \tag{9}$$

so that the domain of definition D_ϵ of the transformed Hamiltonian is the set of twice differentiable functions $f(\theta)$ (with $\theta \in [0, 2\pi]$), such that

$$f(0) = e^{i2\pi\epsilon} f(2\pi), \tag{10}$$

$$\frac{df}{d\theta}(0) = e^{i2\pi\epsilon}\frac{df}{d\theta}(2\pi). \tag{11}$$

In D_ϵ the Hamiltonian H_g is again essentially selfadjoint and, obviously, has the same eigenvalues as H with the eigenvectors

$$\Psi_{g,n}(\epsilon, \theta) = \frac{1}{\sqrt{2\pi}}e^{i(n-\epsilon)\theta}. \tag{12}$$

In this gauge, the kinetic part of the Hamiltonian H_g does not depend on ϵ whereas its eigenvectors and the domain of definition depend and, consequently, it is clear that the HF equation (1) is not valid for H_g since

$$\left\langle \Psi_{g,n} \left| \frac{\partial H_g}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) \right. \right\rangle = 0. \tag{13}$$

Now, the HF theorem fails because with the gauge transformation we have transferred the dependence on the electromagnetic field from the kinetic part of H to the domain of definition of H_g, a situation that is not taken into account in this form of the theorem. The key point here is that, because D_ϵ depends on ϵ, some quantities like

$$\lim_{t\to 0}\langle \Psi_{g,n} | \frac{H_g(\epsilon + t) - H_g(\epsilon)}{t} | \Psi_{g,n}\rangle \tag{14}$$

must be handled carefully because the domains of definition of $H_g(\epsilon + t)$ and $H_g(\epsilon)$ are different and H_g is a non-bounded operator.

3. Generalization of the Hellmann–Feynman theorem

In the previous section, we have seen an example that illustrates the need of a generalization of HF theorem in order to include the cases in which the domain of definition of the Hamiltonian also depends on the parameter we differentiate with respect to.

Consider a selfadjoint Hamiltonian $H(\lambda)$ with domain $\mathcal{D}_\lambda$, eigenvalues $E_n(\lambda)$ and normalized eigenvectors $\Psi_n(\lambda)$. Assume differentiability with respect to the parameter λ.

Taking derivatives in the identity

$$E_n(\lambda) = \left\langle \Psi_n(\lambda) | H(\lambda)\Psi_n(\lambda)\right\rangle \tag{15}$$

we find

$$\frac{\partial E_n}{\partial \lambda} = \left\langle \frac{\partial \Psi_n}{\partial \lambda} \middle| H\Psi_n \right\rangle + \left\langle \Psi_n \middle| \frac{\partial H}{\partial \lambda}\Psi_n \right\rangle + \left\langle \Psi_n \middle| H\frac{\partial \Psi_n}{\partial \lambda} \right\rangle.$$

Note that in order to make sense of the previous expression, one has to extend $H(\lambda)$ to an operator defined in a domain that includes $\mathcal{D}_\lambda$ for any λ. In this case, the extended Hamiltonian will not be selfadjoint in general and therefore the expressions that involve such a non-Hermitian operator must be handled with care. That is the key point of the rest of the Letter.

Adding and subtracting the term $\langle H\Psi_n | \partial_\lambda \Psi_n \rangle = E_n\langle \Psi_n | \partial_\lambda \Psi_n \rangle$, and using the fact that

$$\left\langle \frac{\partial}{\partial \lambda}\Psi_n(\lambda) \middle| \Psi_n(\lambda) \right\rangle + \left\langle \Psi_n(\lambda) \middle| \frac{\partial}{\partial \lambda}\Psi_n(\lambda) \right\rangle$$

$$= \frac{\partial}{\partial \lambda}\langle \Psi_n(\lambda) | \Psi_n(\lambda)\rangle = 0 \tag{16}$$

for normalized eigenvectors, we finally obtain the desired generalization of the Hellmann–Feynman theorem:

$$\frac{\partial E_n(\lambda)}{\partial \lambda} = \left\langle \Psi_n(\lambda) \middle| \frac{\partial H(\lambda)}{\partial \lambda} \middle| \Psi_n(\lambda)\right\rangle + \Delta_n(\lambda) \tag{17}$$

with

$$\Delta_n(\lambda) = \left\langle \Psi_n(\lambda) \left| H(\lambda) \frac{\partial}{\partial \lambda} \Psi_n(\lambda) \right. \right\rangle - \left\langle H(\lambda) \Psi_n(\lambda) \left| \frac{\partial}{\partial \lambda} \Psi_n(\lambda) \right. \right\rangle. \quad (18)$$

Compared with Eq. (1), we have a new term, Δ_n, that vanishes for states $\Psi_n(\lambda)$ in $\mathcal{D}_\lambda$ such that $\partial_\lambda \Psi_n(\lambda)$ is also in $\mathcal{D}_\lambda$. In this case the normal form of HF theorem (1) is recovered. However, in general, for states $\Psi_n(\lambda)$ such that $\partial_\lambda \Psi_n(\lambda) \notin \mathcal{D}_\lambda$, the new factor $\Delta_n(\lambda)$ can give an extra contribution to the second term of the HF equation (17).

The generalization of the other forms of the HF theorem can be obtained in a similar way. For the integrated form we have

$$\big(E_n(\lambda_1) - E_n(\lambda_2)\big) \big\langle \Psi_n(\lambda_2) \big| \Psi_n(\lambda_1) \big\rangle$$
$$= \big\langle \Psi_n(\lambda_2) \big| \big(H(\lambda_1) - H(\lambda_2)\big) \Psi_n(\lambda_1) \big\rangle + \Delta_n(\lambda_1, \lambda_2), \quad (19)$$

with

$$\Delta_n(\lambda_1, \lambda_2)$$
$$= \big\langle \Psi_n(\lambda_2) \big| H(\lambda_2) \Psi_n(\lambda_1) \big\rangle - \big\langle H(\lambda_2) \Psi_n(\lambda_2) \big| \Psi_n(\lambda_1) \big\rangle. \quad (20)$$

And the off diagonal formulation is

$$\big(E_m(\lambda) - E_n(\lambda)\big) \left\langle \Psi_n(\lambda) \left| \frac{\partial}{\partial \lambda} \Psi_m(\lambda) \right. \right\rangle$$
$$= \left\langle \Psi_n(\lambda) \left| \frac{\partial H(\lambda)}{\partial \lambda} \Psi_m(\lambda) \right. \right\rangle + \Delta_{n,m}(\lambda) \quad (21)$$

where the anomalous term is now

$$\Delta_{n,m}(\lambda) = \left\langle \Psi_n(\lambda) \left| H(\lambda) \frac{\partial \Psi_m(\lambda)}{\partial \lambda} \right. \right\rangle - \left\langle H(\lambda) \Psi_n(\lambda) \left| \frac{\partial \Psi_m(\lambda)}{\partial \lambda} \right. \right\rangle. \quad (22)$$

4. Applications

Now we can apply Eqs. (17), (18) to the charged planar rotor interacting with a confined magnetic field that was analyzed in Section 2. For the Hamiltonian H defined in (4) we obtain $\Delta_n(\epsilon) = 0$ (because in this case $\partial_\epsilon \Psi_n \in D_0$) and we recover Eq. (7). However, for H_g we have:

$$\frac{\partial}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) = -i\theta \Psi_{g,n}(\epsilon, \theta). \quad (23)$$

Consequently $\partial_\epsilon \Psi_{g,n}(\epsilon, \theta)$ is not in D_ϵ. Now we can evaluate Δ_n as:

$$\Delta_n = \left\langle \Psi_{g,n}(\epsilon, \theta) \left| H_g \frac{\partial}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) \right. \right\rangle$$
$$- \left\langle H_g \Psi_{g,n}(\epsilon, \theta) \left| \frac{\partial}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) \right. \right\rangle, \quad (24)$$

that integrating by parts (or evaluated directly) gives a boundary term:

$$\Delta_n = \frac{1}{2} \left\{ \left[\frac{\partial}{\partial \theta} \Psi_{g,n}^*(\epsilon, \theta) \frac{\partial}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) \right]_0^{2\pi} \right.$$
$$\left. - \left[\Psi_{g,n}^*(\epsilon, \theta) \frac{\partial}{\partial \theta} \frac{\partial}{\partial \epsilon} \Psi_{g,n}(\epsilon, \theta) \right]_0^{2\pi} \right\}$$
$$= \epsilon - n. \quad (25)$$

It is interesting to note that in this case, the origin of the extra term Δ_n can be related to the anomalous Virial theorem for this kind of systems [6].

Another example of a system where the standard form (1) of the HF theorem fails is a nonrelativistic particle in two dimensions interacting with a $\delta(r)/r$ potential. This system has been studied in

[7,8] as an example of anomalous symmetry breaking in quantum mechanics. The Hamiltonian

$$H = -\frac{\hbar^2}{2m} \nabla^2 + \lambda \delta^2(\mathbf{r}) \quad (26)$$

transforms under dilations, $\mathbf{r} \mapsto \alpha \mathbf{r}$, in a homogeneous way $H \mapsto \alpha^{-2} H$. This means that it cannot have any normalizable eigenvector with energy different from zero, since if $\Psi_n(\mathbf{r})$ is an eigenstate with energy E_n, then $\Psi_n(\alpha \mathbf{r})$ is also an eigenstate with energy $\alpha^{-2} E_n$ for any real number α. If H is selfadjoint the previous property implies that the only point in the discrete spectrum is $E = 0$.

It is also known, however, that in this system the $SO(2, 1)$ symmetry is anomalously broken and there is one s-wave normalized eigenstate

$$\Psi_0(\alpha, \mathbf{r}) = \frac{\alpha}{\pi^{1/2}} K_0(\alpha r), \quad (27)$$

which has the energy

$$E_0 = -\frac{\hbar^2}{2m} \alpha^2, \quad (28)$$

where in (27) the K_0 Bessel function was used and the value of α determines the domain of the Hamiltonian. The apparent contradiction is solved by the fact that although the Hamiltonian transforms multiplicatively by a dilation its domain is not preserved and therefore the dilation does not actually produce any new eigenvectors for the original Hamiltonian.

To see that, we can solve the s-wave sector of the Hamiltonian (26) by renormalizing the coupling λ as in [7] or by working in $\mathbb{R}^2 \setminus \{(0,0)\}$, in order to avoid the singularity at the origin [8,9]. In polar coordinates with Hamiltonian

$$H_s = -\frac{\hbar^2}{2m} \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} \right), \quad (29)$$

defined on the domain

$$D_0^\beta = \left\{ f \in L^2(R_+, r\,dr) \,\middle|\, f \in C^\infty(R_+),\ c_0 = \beta c_1 \right\} \quad (30)$$

with

$$c_0 = \lim_{r \to 0} \left(\frac{f(r)}{\log(\alpha_0 r)} \right), \quad (31)$$

$$c_1 = \lim_{r \to 0} \big[f(r) - c_0 \log(\alpha_0 r) \big]. \quad (32)$$

In this domain H_s is essentially selfadjoint. The meaning of this conditions is that if $f(r) \in D_0^\beta$ then for $r \to 0$; $f(r) \sim a(\log(\alpha r) + b) + o(1)$ with

$$\frac{1}{\beta} = b + \log\left(\frac{\alpha}{\alpha_0} \right) \quad (33)$$

and α_0 the subtraction point. Now, it is easy to see that $\partial_\alpha \Psi_0(\alpha, \mathbf{r}) = (2/\alpha) G \Psi_0(\alpha, \mathbf{r})$ where $G = \frac{1}{4}(\mathbf{xp} + \mathbf{px})$ is the infinitesimal generator of the dilations symmetry which is anomalously broken because if $\Psi_0(\alpha, \mathbf{r}) \in D_0^\beta$ then $G \Psi_0(\alpha, \mathbf{r}) \in D_0^{\beta'}$ with $\beta' = \frac{\beta}{\beta+1}$. As the differentiation with respect to the parameter α does not preserve the domain, only the general form of the HF theorem is valid.

In fact one can compute

$$\frac{dE_0}{d\alpha} = -\frac{\hbar^2}{m} \alpha, \quad (34)$$

$$\frac{dH_s}{d\alpha} = 0, \quad (35)$$

$$\Delta_0 = \frac{\hbar^2 \alpha}{m} \left\{ \left[r \frac{d}{dr} K_0(\alpha r) \left(r \frac{d}{dr} + 1 \right) K_0(\alpha r) \right. \right.$$

$$\left. - rK_0(\alpha r)\frac{d}{dr}\left(r\frac{d}{dr} + 1\right)K_0(\alpha r)\right]_0^{\infty}\right\}$$
$$= -\frac{\hbar^2 \alpha}{m} \tag{36}$$

and the generalized Hellman–Feynman theorem (17) holds. Notice that the evaluation of Δ_0 does not require the knowledge of Ψ_0 but only its behaviour at $r = 0$ and $r = \infty$, which is fixed by the boundary conditions on the domain D_0^{β}.

There is a similarity between this system and the one analyzed previously, here we have eliminated the divergent term of the potential by avoiding the origin and introducing some boundary conditions for the functions of the domain of the Hamiltonian in that point and, in this way, we have transferred the dependence with the parameter from the Hamiltonian to its domain of definition, exactly the same that happened in the first example. On the other hand, the contribution of the Δ_0 term is related to the property that G does not keep the domain of H invariant, a fact whose main consequence is the appearance of a conformal anomaly in the system.

5. Conclusion

The main objective of this Letter is to present a generalization of the well-known Hellmann–Feynman theorem to enlarge its range of applicability. In particular, our generalization allows to deal with several interesting cases where the Hamiltonian of the system depends on a parameter, not only in the explicit expression of the operator, but also in its domain of definition. We also present some remarkable examples to show a few applications of the generalization of the theorem.

Acknowledgements

References

[1] R.P. Feynman, Phys. Rev. 56 (1939) 340.
[2] P. Bala, B. Lesyng, J.A. McCammon, Chem. Phys. Lett. 219 (1994) 259.
[3] F.M. Fernández, Chem. Phys. Lett. 233 (1995) 651.
[4] P. Ziesche, K. Kunze, B. Milek, J. Phys. A: Math. Gen. 20 (1987) 2859.
[5] M. Asorey, J.G. Esteve, A.F. Pacheco, Phys. Rev. D 27 (1983) 1852.
[6] J. Abad, J.G. Esteve, Phys. Rev. A 44 (1991) 4728.
[7] R. Jackiw, in: A. Ali, P. Hoodbhoy (Eds.), M.A. Beg Memorial Volume, World Scientific, Singapore, 1991, pp. 25–42.
[8] J.G. Esteve, Phys. Rev. D 66 (2002) 125013.
[9] S. Albeverio, K. Gesztesy, R. Høegh-Krohn, H. Holden, Solvable Models in Quantum Mechanics, Springer, Heidelberg, 1988.