Degenerate Ground States and a Fractional Number of Electrons in Density and Reduced Density Matrix Functional Theory

Weitao Yang and Yingkai Zhang
Department of Chemistry, Duke University, Durham, North Carolina 27708

Paul W. Ayers
Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina 27599

(Received 26 August 1999)

For a linear combination of electron densities of degenerate ground states, it is shown that the value of any energy functional is the ground state energy, if the energy functional is exact for ground state densities, size consistent, and translational invariant. The corresponding functional of kinetic and interaction energy is the linear combination of the functionals of the degenerate densities. Without invoking ensembles, it is shown that the energy functional of fractional number electrons is a series of straight lines interpolating its values at integers. These results underscore the importance of grand canonical ensemble formulation in density functional theory.

PACS numbers: 71.15.Mb, 31.15.Ew

Density functional theory (DFT) [1–4] is a rigorous approach for describing the ground state of any electronic system. Much progress has been made in the theoretical formulation, approximation, and application of functionals. The success of DFT is based on the quality of approximation for the energy functional. The construction of accurate approximations depends on the theoretical formulation.

The original Hohenberg-Kohn theorem only formulates the energy functional for v-representable densities—densities which can be constructed from a ground state wave function (which are also called pure-state v-representable densities). Since there is no easy way to tell when a given density is v representable, it is necessary to extend the domain of the energy functional to an easily defined set of densities. Towards this goal, Levy [3,5] and Lieb [6] extended the energy functional to all nonnegative, continuous densities which integrate to an integer number of electrons. Such densities are said to be N representable, as they can be constructed from a pure-state many-electron wave function. The extension to densities which correspond to mixed states was made by Valone [7]. Since DFT uses the electron density as the fundamental variable, it is also necessary to treat systems with fractional numbers of electrons [8–10]. Such an extension was made by Perdew, Parr, Levy, and Balduz [11].

In this Letter, we examine these two types of densities, non-v-representable densities and densities which integrate to a fractional number of electrons. We show that such densities can arise as the densities of isolated subsystems inside a "supermolecular" system which is described by a ground state wave function. We prove that density functionals which are exact for v-representable densities, size consistent, and translationally invariant behave like density functionals formulated through constrained search in the grand canonical ensemble. This underscores the importance of the grand canonical ensemble-based formulation of DFT. Similar results are also obtained for energy functionals of first-order reduced density matrices (1-RDM).

How non-v-representable densities arise as subsystem densities for isolated subsystems inside a supermolecular system is conveniently illustrated by a simple example. Consider the case of N electrons in an external potential $v_{\mathbf{R}}(\mathbf{r})$ which has a threefold degeneracy in its ground state with the ground state energy $E^0_{v_{\mathbf{R}}}(N)$. We use $\mathbf{R}$ to indicate the location of the external potential. For example, we can just take $\mathbf{R}$ to be the center-of-mass vector of the nuclei generating the external potential. The three-dimensional vector space of the ground state can be described by three *orthogonal and degenerate* ground state wave functions, $\Phi_A(\mathbf{R})$, $\Phi_B(\mathbf{R})$, and $\Phi_C(\mathbf{R})$, where the coordinates of the electrons are not written explicitly. The Schrödinger equations for Φ_A is

$$\hat{H}_{\mathbf{R}}\Phi_A(\mathbf{R}) = E^0_{v_{\mathbf{R}}}(N)\Phi_A(\mathbf{R}), \tag{1}$$

where the Hamiltonian $\hat{H}_{\mathbf{R}}$ is composed of the N-electron kinetic energy, electron-electron interaction energy, and the potential term from $v_{\mathbf{R}}(\mathbf{r})$. Similar equations follow for Φ_B and Φ_C.

Now construct a $3N$-electron system in the external potential $v(\mathbf{r}) = v_{\mathbf{R}_1}(\mathbf{r}) + v_{\mathbf{R}_2}(\mathbf{r}) + v_{\mathbf{R}_3}(\mathbf{r})$. The Hamiltonian is $\hat{H} = \hat{H}_{\mathbf{R}_1} + \hat{H}_{\mathbf{R}_2} + \hat{H}_{\mathbf{R}_3}$. We consider our system only at the *infinite-separation limit*; namely, $|\mathbf{R}_l - \mathbf{R}_{l'}| \rightarrow \infty$, for all $l \neq l'$. Then the ground state of the system is simply composed of three isolated N-electron subsystem ground states with three identical external potentials separated by infinite distances, assuming the convexity condition

$$E^0_{v_{\mathbf{R}}}(N) \leq [E^0_{v_{\mathbf{R}}}(N+1) + E^0_{v_{\mathbf{R}}}(N-1)]/2, \tag{2}$$

which is known to hold for atoms and molecules from experimental data [4,11]. The wave function is just the antisymmetric product of three separate ones. Since each of the three N-electron systems has threefold degeneracy in the ground state, described by Φ_A, Φ_B, and Φ_C, we thus have a total of 27-fold degeneracy for the ground state of our $3N$-electron system. One of the degenerate ground state wave functions is

$$\Psi_1 = \hat{A}(\Phi_A(\mathbf{R}_1)\Phi_B(\mathbf{R}_2)\Phi_C(\mathbf{R}_3)), \tag{3}$$

where $\hat{A}$ is the $3N$-electron antisymmetrization operator. There are a total of six (3!) wave functions $[\Psi_k, k = 1, \ldots, 6]$ of this type—each has three different N-electron wave functions at the three different locations. Note that $\Phi_A(\mathbf{R}_1)$, $\boldsymbol{\Phi}_\mathbf{B}(\mathbf{R}_2)$, and $\Phi_C(\mathbf{R}_3)$ are ground state wave functions for $v_{\mathbf{R}_1}$, $v_{\mathbf{R}_2}$, and $v_{\mathbf{R}_3}$, respectively. The energy of any of the wave functions is $3E^0_{v_\mathbf{R}}(N)$.

Clearly, any linear combination of the $\{\Psi_k, k = 1, \ldots, 6\}$ is also a ground state wave function of our system with energy $3E^0_{v_\mathbf{R}}(N)$. Consider the averaged one; namely,

$$\Psi = \left\{\sum_{k=1}^{6}\Psi_k\right\}\bigg/\sqrt{6}\,. \tag{4}$$

The electron density of this wave function is

$$\rho_{123} = \sum_{l=1}^{3}\rho^l = \sum_{l=1}^{3}\frac{1}{3}(\rho_A^l + \rho_B^l + \rho_C^l), \tag{5}$$

where ρ_A^l, for example, denotes the electron density of $\Phi_A(\mathbf{R}_l)$. We note that the cross terms like $N\langle\Psi_k|\delta(\mathbf{r} - \mathbf{r}_1)|\Psi_{k'}\rangle$ vanish because Φ_A and Φ_B are orthogonal and the subsystems are infinitely separated. The density of Eq. (5) is simply composed of three identical densities $\rho^l = \frac{1}{3}(\rho_A^l + \rho_B^l + \rho_C^l)$, $l = 1, 2, 3$, each separated from the other by an infinity distance. The energy of this wave function and density is $3E^0_{v_\mathbf{R}}(N)$. *While the density ρ^l is non-v-representable [5,6], it is the density of an isolated subsystem inside the supermolecule described by the wave function of Eq. (4).* This reveals a surprising nonlocal property of quantum mechanics.

Now we consider the behavior of the energy functional, $E_v[\rho]$, for this density. We do not restrict ourselves to any specific formulation of $E_v[\rho]$, instead assuming that $E_v[\rho]$ possesses certain useful properties. Namely:

(1) $E_v[\rho]$ is exact for any (pure-state) v-representable densities. Hence, for the total density in Eq. (5), we have

$$E_v[\rho_{123}(\mathbf{r})] = 3E^0_{v_\mathbf{R}}(N)\,. \tag{6}$$

(2) $E_v[\rho]$ is size consistent. Therefore,

$$E_v[\rho_{123}(\mathbf{r})] = \sum_{l=1}^{3} E_{v_{\mathbf{R}_l}}\left[\frac{1}{3}(\rho_A^l + \rho_B^l + \rho_C^l)\right]. \tag{7}$$

(3) $E_v[\rho]$ is translationally invariant. Therefore,

$$E_{v_{\mathbf{R}_1}}[\rho^1] = E_{v_{\mathbf{R}_2}}[\rho^2] = E_{v_{\mathbf{R}_3}}[\rho^3]. \tag{8}$$

From Eqs. (6)–(8) it follows that

$$E_{v_{\mathbf{R}_l}}\left[\frac{1}{3}(\rho_A^l + \rho_B^l + \rho_C^l)\right] = E^0_{v_\mathbf{R}}(N)\,. \tag{9}$$

This shows that for *any* energy functional which possesses the desirable properties of exactness, size consistency, and translational invariance, its value for the non-v-representable density $\frac{1}{3}(\rho_A^l + \rho_B^l + \rho_C^l)$ is just the corresponding ground state energy, $E^0_{v_\mathbf{R}}(N)$.

Note that here we define the *functional size-consistency* condition with Eq. (7). There is also an *energy size-consistency* condition for any quantum mechanical method defined by the ground state energy expression $\tilde{E}^0_v(N)$ for N electrons in an external potential v; it is, in our example, $\tilde{E}^0_v(3N) = \sum_{l=1}^{3}\tilde{E}^0_{v_{\mathbf{R}_l}}(N)$. The functional size-consistency condition leads to the energy size-consistency condition, but the converse is not always true.

We now generalize our result of Eq. (9) to include arbitrary linear combinations of degenerate densities, and arbitrary numbers of degenerate densities. The final result is the following: For an N-electronic system in potential $v(\mathbf{r})$ that has g degenerate ground state wave functions $(\Phi_i, i = 1, 2, \ldots, g)$ with corresponding densities $(\rho_i, i = 1, 2, \ldots, g)$ and ground state energy $E^0_v(N)$, we can construct the density $\rho = \sum_{i=1}^{g} C_i\rho_i$, where $\{C_i\}$ are positive fractional numbers, and satisfy the normalization condition $\sum_{i=1}^{g} C_i = 1$. The exact energy functional should satisfy the following equation

$$E_v\left[\sum_{i=1}^{g} C_i\rho_i\right] = E_v[\rho_i] = E^0_v(N)\,. \tag{10}$$

The proof of Eq. (10) is very similar to that of Eq. (9). We here provide the proof for the case of $g = 2$ with the two degenerate wave functions Φ_A of Eqs. (1) and Φ_B. Consider a qN-electron system in the external potential $v(\mathbf{r}) = \sum_{l=1}^{q} v_{\mathbf{R}_l}(\mathbf{r})$ at the infinite-separation limit. Then the system is simply composed of q noninteracting N-electron systems with q identical external potentials separated by infinite distances, assuming the convexity condition of Eq. (2). The total ground state wave function is the antisymmetric product of q separated N-electron ground state wave functions, with a total energy of $qE^0_{v_\mathbf{R}}$. One such possibility is for the first p locations, $\mathbf{R}_1 \cdots \mathbf{R}_p$, to have the wave function Φ_A and the rest the other wave function Φ_B; namely,

$$\Psi_1 = \hat{A}\{\Phi_A(\mathbf{R}_1)\Phi_A(\mathbf{R}_2)\cdots\Phi_A(\mathbf{R}_p) \tag{11}$$

$$\Phi_B(\mathbf{R}_{p+1})\Phi_B(\mathbf{R}_{p+2})\cdots\Phi_B(\mathbf{R}_q)\}\,. \tag{12}$$

Permutation of any two locations with Φ_A and Φ_B generates a different degenerate qN-electron wave function. There are a total of $m = \frac{q!}{p!(q-p)!}$ such degenerate wave functions, $\{\Psi_1, \Psi_2 \ldots \Psi_m\}$. For any wave function Ψ_k, a particular location $\mathbf{R}_s$ can either have the wave function Φ_A or Φ_B. In all such wave functions $\{\Psi_1, \Psi_2 \ldots \Psi_m\}$, the number of times any location $\mathbf{R}_s$ having the wave function Φ_A is equal to $m_A = \frac{(q-1)!}{(p-1)!(q-p)!}$, and the corresponding number for Φ_B is $m_B = m - m_A$. In analogy to Eq. (4), the following qN-electron wave function is also a degenerate wave function

$$\Psi = \frac{1}{\sqrt{m}} \sum_{k=1}^{m} \Psi_k \, , \tag{13}$$

the density of which is

$$\begin{aligned}\rho &= \frac{1}{m} \left\{ \sum_{l=1}^{q} [m_A \rho_A^l + (m - m_A) \rho_B^l] \right\} \\ &= \sum_{s=1}^{q} \left(\frac{p}{q} \rho_A^l + \frac{q - p}{q} \rho_B^l \right) .\end{aligned} \tag{14}$$

Following the arguments leading to Eqs. (6)–(9), we have

$$E_{v_{\mathbf{R}_l}} \left[\left(\frac{p}{q} \, \rho_A^l + \frac{q - p}{q} \, \rho_B^l \right) \right] = E^0_{v_{\mathbf{R}}}(N) \, , \tag{15}$$

which is just Eq. (10) for the case of $g = 2$. Our proof can also be extended to arbitrary number of g, but we skip the details here.

We now explore the consequence of Eq. (10). In density functional theory, the energy functional is normally expressed as

$$E_v[\rho] = F[\rho] + \int \rho v \, d\mathbf{r} \, , \tag{16}$$

where $F[\rho]$ is a universal density functional. From Eq. (10) it follows $E_v[\sum_{i=1}^{g} c_i \rho_i] = \sum_{i=1}^{g} c_i E_v[\rho_i]$, using the fact $\sum_{i=1}^{g} c_i = 1$. This and Eq. (16) give

$$F\left[\sum_{i=1}^{g} c_i \rho_i \right] = \sum_{i=1}^{g} c_i F[\rho_i] \, . \tag{17}$$

We now follow the usual decomposition, $F[\rho] = T_s[\rho] + E_{\mathrm{xc}}[\rho] + J[\rho, \rho]$, where $T_s[\rho]$ is the noninteracting Kohn-Sham kinetic energy, $E_{\mathrm{xc}}[\rho]$ is the exchange correlation energy, and $J[\rho, \rho]$ is the Coulomb interaction $J[\rho_1, \rho_2] = \frac{1}{2} \int \int \frac{\rho_1(\mathbf{r}) \rho_2(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \, d\mathbf{r} \, d\mathbf{r}'$. Equation (17) then becomes

$$T_s\left[\sum_{i=1}^{g} c_i \rho_i \right] + E_{\mathrm{xc}}\left[\sum_{i=1}^{g} c_i \rho_i \right] = \sum_{i=1}^{g} c_i (T_s[\rho_i] + E_{xc}[\rho_i]) + \sum_{i,j}^{g} (c_i \delta_{ij} - c_i c_j) J[\rho_i, \rho_j] \, . \tag{18}$$

Equations (10), (17), and (18) place stringent conditions on the density functionals. No known approximate functional satisfies Eq. (17). For any energy functional satisfying Eqs. (6)–(8), according to Eq. (10), for an open-shell atom with g-fold degeneracy, *the energy for the spherically averaged density* $\sum_{i=1}^{g} \frac{1}{g} \rho_i$ *should just be the ground state energy*. This is not satisfied by any explicit functional known so far: In practical DFT calculations for open-shell atoms, it is well known that if one uses a spherically averaged density, the calculated atomic energy is significantly higher than those from nonaveraged densities, and the latter give much more accurate atomic energies.

We now deal with the case of fractional number of electrons with *similar* techniques. Consider a system with $qN + p$ electrons, where $p, q(>p)$ and N are integers, in the external potential $v(\mathbf{r}) = \sum_{l=1}^{q} v_{\mathbf{R}_l}(\mathbf{r})$ at the infinite-separation limit. Then the total system is simply composed of q noninteracting systems with identical external potential separated by infinite distances. For the ground state of this system, the only possible distribution of the number of electrons among the q systems is for p systems to have $N + 1$ electrons and $(q - p)$ systems to have N electrons, assuming the convexity condition of Eq. (2). Let Φ_{N+1} be the ground state wave function for the $(N + 1)$-electron system and Φ_N be the ground state wave function for the N-electron system under the same potential $v_{\mathbf{R}}$. The total ground state wave function is the antisymmetric product of q separated ground state wave functions, with a total energy of $(q - p)E_v^0(N) + pE_v^0(N + 1)$. One such possibility is for the first p locations, $\mathbf{R}_1 \cdots \mathbf{R}_p$, to have the wave function Φ_{N+1} and the rest the other wave function Φ_N; namely,

$$\begin{aligned}\Psi_1 = \hat{A}\{&\Phi_{N+1}(\mathbf{R}_1)\Phi_{N+1}(\mathbf{R}_2) \cdots \Phi_{N+1}(\mathbf{R}_p) \\ &\times \Phi_N(\mathbf{R}_{p+1})\Phi_N(\mathbf{R}_{p+2}) \cdots \Phi_N(\mathbf{R}_q)\} \, .\end{aligned} \tag{19}$$

Permutation of any two locations with Φ_{N+1} and Φ_N generates a different degenerate wave function. There are a total of $m = \frac{q!}{p!(q-p)!}$ such degenerate wave functions, $\{\Psi_1, \Psi_2 \cdots \Psi_m\}$. For any wave function Ψ_k, a particular location $\mathbf{R}_l$ can either have the wave function Φ_N or Φ_{N+1}. In all such wave functions $\{\Psi_1, \Psi_2 \cdots \Psi_m\}$, the number of times any location $\mathbf{R}_l$, having the wave function Φ_{N+1} is equal to $m_{N+1} = \frac{(q-1)!}{(p-1)!(q-p)!}$, and the corresponding number for Φ_N is $m_N = m - m_{N+1}$. In analogy to Eq. (13), $\Psi = (1/\sqrt{m}) \sum_{k=1}^{m} \Psi_k$ is also a degenerate wave function, the density of which is

$$\begin{aligned}\rho &= \frac{1}{m} \left\{ \sum_{l=1}^{q} [m_{N+1} \rho_{N+1}^l + (m - m_{N+1}) \rho_N^l] \right\} \\ &= \sum_{l=1}^{q} \left(\frac{p}{q} \rho_{N+1}^l + \frac{q - p}{q} \rho_N^l \right) .\end{aligned} \tag{20}$$

Following the arguments of Eqs. (6)–(9), we have

$$\begin{aligned}&E_v\left[\frac{p}{q} \, \rho_{N+1} + \frac{q - p}{q} \, \rho_N \right] \\ &\qquad = \frac{p}{q} \, E_v^0(N + 1) + \frac{q - p}{q} \, E_v^0(N) \, ,\end{aligned} \tag{21}$$

where we have dropped the reference to the location $\mathbf{R}_l$.

Equation (21) *defines* the ground state energy of fractional number of electrons as

$$E_v^0\left(N + \frac{p}{q}\right) = \frac{p}{q} E_v^0(N + 1) + \frac{q - p}{q} E_v^0(N), \tag{22}$$

which is identical to the result of Ref. [11]. However, the present derivation does not invoke the grand canonical ensemble, instead using the conditions of "exact for v-representable densities," "size consistency," and "translational invariance" of the energy density functional.

Note that Savin has also emphasized the importance of the grand canonical ensemble formulation of DFT [12] and Perdew [8] has showed that the expectation value of the wave function for a system with fractional number of electrons is the same as the ensemble expectation value. Recently Schipper *et al.* have presented two molecular examples concerning v representability [13].

Equations (10) and (21) are also valid for energy functionals of one-dimensional reduced density matrix (1-RDM) γ because Eqs. (14) and (20) are valid for the corresponding 1-RDM. In particular, Eq. (21) becomes

$$E_v\left[\frac{p}{q}\gamma_{N+1} + \frac{q - p}{q}\gamma_N\right] = \frac{p}{q} E_v^0(N + 1) + \frac{q - p}{q} E_v^0(N). \tag{23}$$

We can extend the canonical ensemble constrained-search formulation of the 1-RDM functional [3,7] to the following grand canonical ensemble formulation:

$$E_v[\gamma] = \min_{\hat{\Gamma}\to\gamma} \mathrm{Tr}[\hat{H}\hat{\Gamma}], \tag{24}$$

where $\hat{H}$ is the Hamiltonian and the minimum is over all γ-constrained density matrices $\hat{\Gamma}$ in the Fock space. Then, the 1-RDM functional satisfies Eq. (23).

In summary, for any size-consistent and translation-invariant energy functional satisfying Eqs. (6)–(8), we proved Eqs. (10), (17), (21), and (23). As a special case of Eq. (10), for the spherically averaged density of an open-shell atom, we also proved that the energy of such a functional is the corresponding ground state energy. Functionals defined only for densities associated with a wave function would have a great deal of difficulty satisfying Eqs. (10) and (17), and functionals defined only for integer number of electrons cannot obey Eqs. (21) or (24). *Only the grand canonical ensemble formulation,* defined in Ref. [11] for the density and defined in Eq. (24) for the 1-RDM, can satisfy Eqs. (17), (21), and (23). Therefore, our results underscore the importance of developing functionals based on the grand canonical ensemble approach.

As a final note, we reexamine the zero-temperature grand canonical ensemble theory. The ground state energy for an arbitrary number of electrons n, including fractional numbers and integers, can be written as [4],

$$\varepsilon_v^0(n) = \min_{p(M)}\left[\sum_M p(M) E_v^0(M)\right], \tag{25}$$

where the summation is over all positive integers, M, $\sum_M p(M) = 1$, $\sum_M M p(M) = n$, and $1 \geq p(M) \geq 0$. With this definition, $\varepsilon_v^0(M) \leq E_v^0(M)$, and the convexity condition, Eq. (2), is valid because

$$\varepsilon_v^0(n) = \min_{p(M)}\left[\sum_M p(M) E_v^0(M)\right] \leq [\varepsilon_v^0(n + 1) + \varepsilon_v^0(n - 1)]/2. \tag{26}$$

Hence, in the grand ensemble definition, $\varepsilon_v^0(n)$ is a nonconcave function. Following Ref. [11], $\varepsilon_v^0(n)$, as a function of the fractional number of electrons, is a piecewise linear interpolation between the integer values. However, $\varepsilon_v^0(M) = E_v^0(M)$ only when $E_v^0(M)$ is nonconcave.

This work has been supported by a grant from the National Science Foundation to Duke University. Discussions with Dr. David Mazziotti have been helpful.

[1] P. Hohenberg and W. Kohn, Phys. Rev. B **136**, 864 (1964).
[2] W. Kohn and L. Sham, Phys. Rev. A **140**, 1133 (1965).
[3] M. Levy, Proc. Natl. Acad. Sci. U.S.A. **76**, 6062 (1979).
[4] R. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
[5] M. Levy, Phys. Rev. A **26**, 1200 (1982).
[6] E. H. Lieb, Int. J. Quantum Chem. **24**, 243 (1983).
[7] S. M. Valone, J. Chem. Phys. **73**, 4653 (1980).
[8] J. P. Perdew, in *Density Functional Methods in Physics,* edited by R. M. Dreizler (Plenum Press, New York, 1985), pp. 265–308.
[9] J. P. Perdew and M. Levy, Phys. Rev. B **56**, 16 021 (1997).
[10] Y. Zhang and W. Yang, J. Chem. Phys. **109**, 2604 (1998).
[11] J. P. Perdew, R. G. Parr, M. Levy, and J. L. Balduz, Jr., Phys. Rev. Lett. **49**, 1691 (1982).
[12] A. Savin, in *Recent Developments and Applications of Modern Density Functional Theory,* edited by J. M. Seminario (Elsevier, Amsterdam, 1996), p. 327.
[13] P. R. T. Schipper, O. V. Gritsenko, and E. J. Baerends, Theor. Chem. Acc. **99**, 329 (1998).