

Climbing the ladder of density functional approximations

John P. Perdew

The following article is an edited transcript of the MRS Materials Theory Award presentation by John P. Perdew on November 26, 2012, at the Materials Research Society Fall Meeting in Boston. The Award "recognizes exceptional advances made by materials theory to the fundamental understanding of the structure and behavior of materials."

Kohn–Sham density functional theory is the most widely used method of electronic-structure calculation in materials physics and chemistry because it reduces the many-electron ground-state problem to a computationally tractable self-consistent one-electron problem. Exact in principle for the ground-state energy and electron density, it requires in practice an approximation to the density functional for the exchange-correlation energy. Common approximations fall on the rungs of a ladder, with higher rungs being more complicated to construct and use but potentially more accurate. Each rung of the ladder introduces an additional ingredient to the energy density. From bottom to top, the rungs are (1) the local spin density approximation, (2) the generalized gradient approximation (GGA), (3) the meta-GGA, (4) the hybrid functional, and (5) the generalized random phase approximation. The semi-local rungs (1–3) are important, because they are computationally efficient, they can be constructed non-empirically, they can serve as input to fourth-rung functionals, and the meta-GGA by itself can be accurate for equilibrium properties. Recent and continuing improvements to the meta-GGA are emphasized here.

Introduction

The subject of my talk is the Kohn–Sham density functional theory (DFT).[1–6] This DFT uses auxiliary occupied orbitals. It is the most widely used method of electronic structure calculation in materials science, condensed matter physics, and quantum chemistry. As evidence of this, the 18 most-cited physics papers published in the period from 1981 to 2010 have been identified.[7] Ten of these papers are concerned with DFT: either fundamental theory, development of approximations, or applications and computational methodologies for DFT.

Kohn–Sham density functional theory

Materials consist of atoms, molecules, nanostructures, solids, and surfaces. They are all systems constructed from many interacting electrons and nuclei. Compared to the electrons, the nuclei are heavy and almost classical, and we often treat them as classical particles that obey Newton's laws of motion. On the other hand, the electrons must be described using quantum mechanics. We are usually interested in the measurable ground-state properties of a system, such as the total energy, E, and changes of the total energy due to adding or removing an electron, stretching or compressing a bond, removing an atom, and other processes. We are also interested in the electron density $n(\vec{r})$, which is a function of position $\vec{r}$ in the material, and in the electron spin densities $n_{\uparrow}(\vec{r})$ and $n_{\downarrow}(\vec{r})$. We further want to know the positions of the nuclei in equilibrium and the vibration frequencies for the nuclei. Kohn–Sham DFT can provide us with answers to all of these questions.

Fundamental quantum mechanics provides answers that are more rigorous and precise, but those answers are computationally very difficult to find. Using the direct quantum mechanical approach with correlated wave functions, it is necessary to solve the many-electron Schrödinger equation. The Hamiltonian or quantum-mechanical energy operator in this equation must include the kinetic energy of all the electrons, the interaction of each electron with an external potential (which is actually the Coulomb attraction between that electron and all of the nuclei), the Coulomb repulsion among the nuclei, and the Coulomb repulsion among the electrons. The wave function is therefore a function of the positions of all N electrons and the z-components of the spins of all N electrons, a very complicated object.[8] We are usually more interested in

John P. Perdew, Department of Physics, Temple University; perdew@temple.edu
DOI: 10.1557/mrs.2013.178

© 2013 Materials Research Society

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

a simpler object such as the density (number of electrons per unit volume at each point in space), which we can obtain in principle if not always in practice by squaring the wave function and integrating over all the coordinates but one (the one representing the chosen point in space) and summing over all of the spins.

In the Schrödinger equation, the motion of each electron is coupled to the motion of every other electron through Coulomb repulsion among the electrons. As a result, the computational time increases very rapidly with the number of electrons in the system, so this method is impractical for large molecules or nanostructures or many other real materials. A computationally efficient alternative is the Kohn–Sham DFT, developed in 1965,[1] for which Walter Kohn received the Nobel Prize in Chemistry in 1998.[8] The main idea of the theory is that we can, in principle, find the exact ground-state energy E and the exact ground-state electron density $n(\vec{r})$ by solving an effective one-electron Schrödinger equation:

$$\hat{h}\psi_{\alpha\sigma} = \varepsilon_{\alpha\sigma}\psi_{\alpha\sigma}, \tag{1}$$

$$\hat{h} = -\frac{1}{2}\nabla^2 + v_{ext}(\vec{r}) + \int d^3r' \frac{n(\vec{r}')}{|\vec{r}' - \vec{r}|} + v_{xc}(\vec{r}), \tag{2}$$

$$n(\vec{r}) = \sum_{\substack{\alpha\sigma \\ \varepsilon_{\alpha\sigma} < \mu}} |\psi_{\alpha\sigma}(\vec{r})|^2, \text{ and}$$

$$v_{xc}(\vec{r}) = \frac{\delta E_{xc}}{\delta n(\vec{r})}.$$

Here $\vec{r}$ and $\vec{r}'$ are different positions in space.

It is much easier to deal with a wave function for one electron (an orbital) than to deal with a correlated wave function for N electrons when N is large. Equation 1 is a one-electron Schrödinger equation with a one-electron Hamiltonian (Equation 2), which is actually an effective Hamiltonian because it includes (in order of presentation) the kinetic energy of that electron, the interaction of that electron with the external potential from the nuclei, the Coulomb repulsion of the electron density acting on that electron, and the exchange-correlation potential, which is only a function of position but in practice has to be approximated. Similarly, the total energy consists of a similar series of terms:

$$E = \sum_{\substack{\alpha\sigma \\ \varepsilon_{\alpha\sigma} < \mu}} \left\langle \psi_{\alpha\sigma} \left| -\frac{1}{2}\nabla^2 \right| \psi_{\alpha\sigma} \right\rangle + \int d^3r n(\vec{r}) v_{ext}(\vec{r}) + \frac{1}{2}\sum_{l \neq l'} \sum_{l'} \frac{Z_l Z_{l'}}{|\vec{R}_l - \vec{R}_{l'}|} + \frac{1}{2}\int d^3r \int d^3r' \frac{n(\vec{r})n(\vec{r}')}{|\vec{r}' - \vec{r}|} + E_{xc}[n]. \tag{3}$$

Equation 3 includes (in order of presentation) the kinetic energy for all of the electrons (obtained from the expectation value of the kinetic energy operator using each occupied orbital), the interaction of the electron density with the external potential, the Coulomb repulsion among the nuclei, and Coulomb repulsion among the electrons. The last term is the exchange-correlation energy E_{xc} as a function of the electron density $n(\vec{r})$; $E_{xc}[n]$ must also be approximated in this theory. The two quantities that must be approximated are related; we obtain the exchange-correlation potential by taking the functional derivative of the exchange-correlation energy with respect to the density. So, in practice, we only need to approximate E_{xc}. If we knew the exact E_{xc} as a functional of the density, we could predict the exact ground-state energy and the exact ground-state density by solving these Kohn-Sham equations self-consistently. In practice, the exact E_{xc} cannot be computed, and we need to make approximations. The way we make these approximations is the main subject of this talk.

A functional is a rule that assigns a number to a function. In this case, it is a rule that assigns an energy, a scalar number, to each electron density $n(\vec{r})$. The exchange-correlation energy is a negative energy that represents the lowering of the energy of the system due to the fact that the electrons avoid each other as they move through the density. The electrons are like a swarm of bees. There is an average density of bees, which is a function of position, and the bees move around within this swarm, but they do not move completely randomly because they sense each other, and they avoid each other. Electrons do the same. Because of the Pauli principle and Coulomb repulsion, electrons avoid each other. This lowers their electrostatic energy, and this energy reduction is the exchange-correlation energy (apart from a small positive piece that represents the kinetic energy of correlation). It is a relatively small fraction of the total energy, but it is very important for all aspects of materials science, because the exchange-correlation energy is what I refer to as "nature's glue."[9] It binds the atoms together into molecules and solids. Without this energy, the atoms would either not be bound together or they would be too loosely bound with weak binding energies and long bond lengths. Without E_{xc}, we would have a completely unrealistic description of materials.

Local density approximation

How do we approximate this exchange-correlation energy functional? The earliest approximation was the local density approximation (LDA) of Kohn and Sham.[1] This is a nonempirical approximation, in the sense that it has no parameter fitted to experimental data. Instead, it uses information about the uniform electron gas, obtained from accurate quantum mechanical calculations, for example from quantum Monte Carlo computations, from which we obtain the exchange-correlation energy per electron of an electron gas of uniform density n. The density is assumed to be uniform over all space. This is a solvable problem. We then treat a nonuniform system—a real atom or molecule or solid—as if it locally consisted of infinitesimal volume elements, each of which is a locally uniform electron gas. This assumption leads to Equation 4:

$$E_{xc}^{LDA}[n] = \int d^3r n(\vec{r}) \varepsilon_{xc}^{unif}(n(\vec{r})), \tag{4}$$

where $\varepsilon_{xc}^{unif}(n)$ is the exchange-correlation energy per particle of an electron gas of uniform density n.[10,11]

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

Kohn and Sham proposed this approximation in 1965. It started to become popular in the early 1970s, when solid-state physicists began to realize that they could obtain accurate results for lattice constants, phonon frequencies, and other properties they wanted to calculate using this approximation. The approximation is exact for a uniform density, and it is correct and accurate for a density that varies slowly over space, but, in fact, it also works reasonably well for atoms, molecules, and surfaces. This was surprising at the time, but the reason was understood in the 1970s:[12,13] The LDA satisfies certain exact constraints on the exchange-correlation energy as a functional of the density. In other words, it includes certain features of the exchange-correlation energy that remain true even when the density is far from uniform. That explanation was the basis for a ladder[14] of approximations of increasing sophistication to this exchange-correlation energy (see **Figure 1**). Today, most chemists and many materials scientists use a higher rung of the ladder than the LDA.

Jacob's ladder of density functional approximations

The strategy behind the construction of the ladder is to try to include additional exact constraints into approximations beyond the LDA, in the hope that those additional exact constraints will make the functional even more accurate. I call this hierarchy "Jacob's ladder of density functional approximations."[14] In the Bible, the patriarch Jacob dreamed of a ladder reaching from earth up to heaven, with the angels ascending and descending. In the Jacob's ladder of density functional approximations, earth is the Hartree approximation ($E_{xc} = 0$), and heaven is the realm of high accuracy. The angels are the users, who can choose the rung that best fits their needs and computational resources. The basic idea is that the exchange-correlation energy is written as an integral over the three-dimensional space of some function $\varepsilon_{xc}^{approx}$ of the local density $n(\vec{r})$ at each point $\vec{r}$, the gradient or derivative of the local density $\nabla n(\vec{r})$ at each point $\vec{r}$, and possibly some other ingredients, such as the orbital kinetic energy density $\tau(\vec{r})$,

$$E_{xc}^{approx}[n] = \int d^3 r n(\vec{r}) \varepsilon_{xc}^{approx}\left(n(\vec{r}), \nabla n(\vec{r}), \tau(\vec{r}), \ldots\right), \qquad (5)$$

where $\tau(\vec{r}) = \frac{1}{2}\sum_{\substack{\alpha\sigma \\ \varepsilon_{\alpha\sigma}<\mu}} \left|\nabla \psi_{\alpha\sigma}(\vec{r})\right|^2$ is the positive kinetic energy density.

The lowest rung of the ladder is the LDA,[1,15] which employs only the local density. The second rung of the ladder is the generalized-gradient approximation (GGA),[16–19] which adds a second ingredient, the gradient of the density. The third rung is the meta-GGA,[20–25] which adds a third ingredient, the orbital kinetic energy density. These first three rungs are semilocal functionals that are computationally very efficient. The fourth and fifth rungs are nonlocal functionals, which are computationally more expensive because they require at least a double integration over three-dimensional space but are potentially more accurate. On the fourth rung, we have hybrid functionals[26–29] that use some exact exchange information, such as the exact exchange energy density or the exact exchange energy itself. On the fifth rung, we have the random-phase-approximation-like functionals[16,30,31] that use the unoccupied Kohn–Sham orbitals, and on that rung DFT starts to resemble traditional many-body theory.

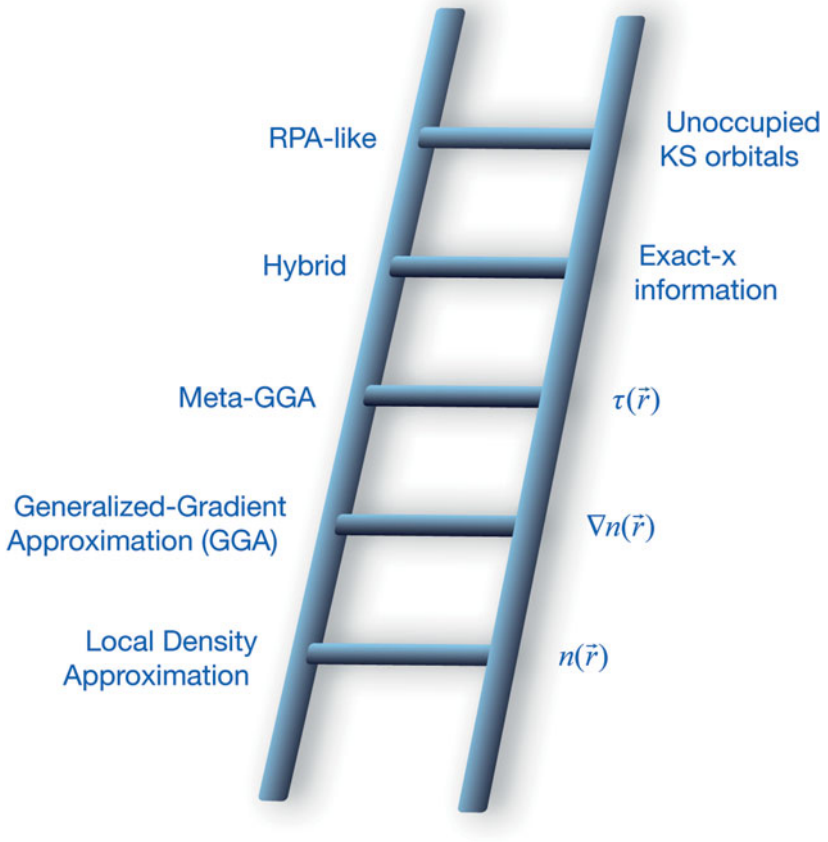


Figure 1. Jacob's ladder of density functional approximations. The rungs are labeled on the left, and their added ingredients are shown on the right.

The orbital kinetic energy density is a positive quantity obtained by summing the squares of the gradients of the occupied orbitals. We can obtain the orbitals from a Kohn–Sham calculation, and then we can easily compute $\tau(\vec{r})$. In the ladder strategy, when we add an ingredient, we can satisfy additional exact constraints, and we can achieve greater accuracy, but for some increase in computational cost.

Some of the exact constraints will now be described. We know a great deal about the exchange-correlation energy, thanks to work by David Langreth, Olle Gunnarson, Bengt

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

Lundqvist, Mel Levy, and others. We also know several exact expressions for E_{xc}. They are not computable in practice, but they are formally exact, for example:[12,13]

$$E_{xc} = \frac{1}{2}\int d^3r n(\vec{r}) \int d^3r' \frac{n_{xc}(\vec{r},\vec{r}')}{|\vec{r}-\vec{r}'|} = E_x + E_c, \quad (6)$$

where $n_{xc}(\vec{r},\vec{r}') = n_x(\vec{r},\vec{r}') + n_c(\vec{r},\vec{r}')$ is the density at $\vec{r}'$ of the (coupling-constant averaged) exchange-correlation hole (*xc* hole) around an electron at $\vec{r}$, with

$$n_x(\vec{r},\vec{r}') \le 0, \int d^3r' n_x(\vec{r},\vec{r}') = -1, \text{and} \int d^3r' n_c(\vec{r},\vec{r}') = 0.$$

Equation 6 tells us that the exchange-correlation energy is the electrostatic interaction between the electron density at point $\vec{r}$ and the density at point $\vec{r}'$ of the exchange-correlation (*xc*) hole that surrounds an electron at $\vec{r}$. This can be divided into an exchange part and a correlation part in a very simple way, and the exact expressions also provide us with certain properties of the hole—exact constraints that can be used to model the functional. For instance, the exchange hole density is non-positive everywhere. The integral of the exchange hole density is –1, and the integral of the correlation hole density is zero. So the idea of the exchange-correlation hole is similar to that of the personal space around a bee in a swarm. There is a region of space around each bee, which the other bees tend to avoid, and this creates a hole in the density. The same thing is true for electrons—each electron is surrounded by a hole in the density, from which we can obtain the exchange-correlation energy.

These constraints help to explain why the LDA works better than expected. The exchange-correlation hole of the LDA is negative, with an integral of –1, and its correlation hole density integrates to zero. These constraints are satisfied automatically by LDA, not only for uniform density but for any density.[12,13] There are other exact constraints that have been discovered over the years, applicable not to the hole but more directly to the energy. If the density varies slowly over space, there is a density gradient expansion for the exchange-correlation energy.[32–34] If the density is scaled uniformly or non-uniformly, there are scaling relations, equalities, and inequalities for the exchange and correlation energies.[35,36] There is a Lieb–Oxford lower bound on the exchange-correlation energy,[37] and there are many other known exact constraints. The higher-level approximations on the higher rungs of the ladder attempt to keep the constraints satisfied by the lower-level rungs, and then add additional constraints in the hope that they will lead to a more accurate functional, and typically they do.

The Hamiltonian in Equation 1 has no magnetic field, which means there is nothing that couples to the *z*-components of the spin or to the current density, and as a result, we only require, in principle, the exchange-correlation energy as a functional of the total density. If we have external fields that couple to the *z*-component of the electron spin, then we need to include the up- and down-spin electron densities; that is spin-DFT.[15] In practice, even if there is no such field coupling to the spin, we use spin-density functionals to achieve higher accuracy at a given level of approximation. It is not difficult to see the reason for this. More information is inputted to a spin-density functional than to a density functional, so the spin-density functional approximations typically work better and provide more useful outputs.

Spin-density functionals are especially useful for open-shell systems, such as typical atoms.[38] We can also use spin-DFT to predict spontaneous magnetization, such as the ferromagnetism of iron. Approximations sometimes predict a spurious spin-symmetry breaking,[13] which leads to a spurious spin density but to a significant improvement in the exchange-correlation energy and in the total energy. This tends to happen in strongly correlated systems. An example of this is stretched H_2. We take the H_2 molecule, stretch out the bond length toward infinity, and try to describe that system through spin-DFT. At some large bond length, the spin symmetry will break in a semi-local approximation, which will no longer yield a singlet ground state but instead will yield a broken-symmetry state with spin up on one nucleus and spin down on the other. This is actually a better description, because it is similar to having separated atoms, so this spin-symmetry breaking can be good for describing the energy at the semi-local level. To simplify the notation, $n_\uparrow$ and $n_\downarrow$ will not be shown, but in practice, all the functionals considered here are spin-density functionals.

Approximation desiderata

What are we looking for when we make an approximation to the exchange-correlation energy? There are three properties that are useful in an approximate functional:

(1) The approximation should be non-empirical. It should be constructed from exact constraints or from a derivation and not from fitting to data sets. The reason is that non-empirical functionals tend to be more universally useful, because the constraints themselves are universal. Non-empirical functionals tend to work for wider classes of systems than functionals fitted to data sets, and they also provide a better test of our understanding of the physics of the exchange-correlation energy. So if a non-empirical functional works, it is probably because we understood something. If an empirical functional works, it works in part because it was fitted. The first three rungs of our ladder, which define the semi-local functionals—local density approximation, GGA, and meta-generalized-gradient approximation—are semi-local; they are not fully non-local, and we have non-empirical constructions for functionals on all of those rungs. On the fourth rung, we have the hybrid functionals, which can be very accurate, but typically they have at least one empirical parameter, and some of them have as many as 40 empirical parameters. The fifth rung includes the random phase approximation (RPA), where we once again have a non-empirical approximation, but we still need to make corrections to the RPA because it is not exact, and, to date, the best corrections

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

are empirical. So the first three rungs are almost completed and non-empirical, and the fourth and fifth rungs still contain some empiricism.

(2) We require this functional to be computationally efficient. When we perform a calculation for a material, we would like to have a result in a week or even in a few days or a few hours. The semi-local functionals are all computationally very efficient, because they evaluate the exchange-correlation energy as a single integral over three-dimensional space, which is computationally simple. The hybrid functionals, which require some exact exchange component, such as the exact exchange energy density,[39,40] can be slower by a factor of 10 or 100, depending on the system and on the numerical program used. The random phase approximation can be 100 times slower than the semi-local functionals. Semi-local functionals require approximately the same computer time. GGA is possibly twice as slow as LDA, while the meta-GGA might require another 30% of the computer time compared to the GGA. That is quite acceptable, because it is not a factor of 10 or 100.

(3) Finally, we require our approximations to be accurate. The approximations need not be exact, but they must provide useful information.

Local density approximation

Let us consider some of these functionals, starting with the LDA[1] or the local spin-density approximation (LSDA):[15]

- It satisfies the exchange-correlation hole sum rules.[12,13]
- It is exact for uniform densities.[10,11]
- It satisfies some of the scaling relations.[35,36]
- It satisfies the general Lieb–Oxford lower bound.[37]
- It is accurate for the linear response of a uniform density.[41]

In practice, the LDA is useful for some simple solids. It predicts lattice constants that are typically 1% or 2% too short, and it predicts total surface energies that are rather realistic. However, if we examine the atomization energy—the energy to break a material or a molecule up into atoms—we find that the LDA yields very serious errors, strongly over-binding atoms into molecules and onto surfaces. So the LDA is not used much in chemistry, but if we are not interested in breaking off atoms, then LDA often works very well.

Generalized gradient approximation

The second rung of the ladder is the GGA,[16–19] which can be constructed to satisfy different sets of exact constraints. There is some incompatibility associated with the limited form of the GGA—we can satisfy some subset of possible constraints, but we cannot satisfy all of them with the same GGA, so we need to select the constraints. One of the standard GGAs is the Perdew–Burke–Ernzerhof (PBE) approximation.[19] PBE recovers the second-order gradient expansion for correlation for a slowly varying density, while preserving the accurate LSDA linear response of a uniform density. It adds a scaling relation for correlation that LSDA does not satisfy. With regard to accuracy, the PBE GGA substantially reduces and improves the overbinding of atoms into molecules and solids. It still overbinds in most cases but much less than LSDA. PBE overcorrects the lattice constants. Calculated lattice constants are too short in LDA, and they are a little too long in GGA. Also, if we examine the surface energy, we find that PBE improves the separate exchange surface energy and correlation surface energy, but the sum of the two is actually less accurate than in LDA. Despite its shortcomings, PBE is still a useful functional.

There is a variant of GGA, "PBEsol"[42] or "PBE for solids," which, unlike PBE, recovers the second-order gradient expansion for the exchange energy over a wide range of slowly varying densities. This is an alternative set of constraints to satisfy; instead of recovering the linear response of the uniform gas, this approximation satisfies the gradient expansion of the exchange energy. PBEsol is in some sense between LDA and PBE. It works very well for simple solids (as do other GGAs for solids)[43,44] and gives accurate lattice constants and accurate surface energies. It can be useful for many solid-state problems, but it is not useful for chemistry problems, because the atomization energies are still too large, somewhere between the strongly overbound LDA atomization energies and the moderately overbound PBE atomization energies. In order to have a really accurate description of typical atoms, molecules, and solids near their equilibrium geometries—an important desideratum—we need to climb to the third rung, the meta-GGA level.

Meta-GGA approximation

Meta-GGAs[20–25] have another ingredient—the orbital kinetic energy density—and we can use that ingredient to satisfy some additional exact constraints. For instance, we can recover the gradient expansion of the exchange energy—not just to the second order as in PBEsol but even to the fourth order[34] in the gradients. At the same time, we can preserve an accurate linear response of the uniform density, which means that the GGA need to select constraints is avoided. The meta-GGAs can also cancel out the correlation energy for any one-electron density, which is desirable because a one-electron system has no correlation. We can also calculate the exchange energy exactly, not for all densities but for some particular density, such as the density of a one-electron atom. By using this extra ingredient—the orbital kinetic energy density—meta-GGA can recognize regions of space in which one orbital shape dominates the electron density. For instance, in the hydrogen and helium atoms, there is only one orbital shape per atom. In a typical covalent chemical bond, there is one dominant orbital shape—the shape of the bonding orbital, and these regions can be treated differently from regions of slowly varying density, even when the density gradient reduces to zero. This means that the meta-GGA is able to recognize the difference between a slowly varying density and a one- or two-electron ground-state density, and this capability has proved to be very useful in correctly describing both solid-state physics and chemistry.

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

One of the meta-GGAs that we favor is revTPSS[23] (a revised Tao–Perdew–Staroverov–Scuseria[22] meta-GGA). In comparison with GGAs, it gives accurate lattice constants for solids simultaneously with accurate surface energies and accurate binding energies for atoms, as well as the binding energies of molecules to surfaces. It performs a number of calculations needed in materials science very well.[44–46] My group recently developed a "meta-GGA made simple,"[24,25] which can even describe intermediate-range van der Waals interactions that are important for the chemistry and physics of soft matter but are less-reliably described in GGAs and in the revTPSS meta-GGA. These semilocal functionals—LDA, GGA, and meta-GGA—will never get the long-range part of the van der Waals interaction right, which requires a nonlocal functional like the RPA. However, they can, in fact, describe the short- and intermediate-range part of the van der Waals interaction, which can be important for binding one atom to another in soft matter. For instance, in graphite, one layer of carbon atoms is bound to the next layer of carbon atoms, not by a covalent bond or a metallic bond, but rather by a van der Waals bond, and much of that van der Waals bond is described by this "meta-GGA made simple." At the same time, it performs like revTPSS for calculating the other properties previously mentioned.[25]

If we have a system that is near mechanical equilibrium, where the bond lengths and the lattice constants are close to their equilibrium values, then a meta-GGA can come very close to satisfying our three desiderata: (1) It is typically non-empirical—there is no need to fit to experimental data sets. (2) It is computationally efficient—not much slower than the LDA. (3) And it has useful accuracy for atoms, molecules, solids, and surfaces for a wide range of systems we might be interested in, provided the bonds are not stretched too much. Meta-GGAs are now available in many standard electronic-structure codes for molecules and solids, including VASP and GAUSSIAN.

We would like to improve the meta-GGA, and we think this is possible because it is a very flexible form. Even for a spin-unpolarized system, it has three arguments. There are rich possibilities for building a better meta-GGA, and we would like to build one that will satisfy all the exact constraints that can be satisfied by a meta-GGA, including one that no one has ever tried to satisfy before, the tight Lieb–Oxford lower bound on the exchange energy of any one-electron density.[37] The exchange energy of any one-electron density is very tightly bounded by its Lieb–Oxford lower bound, which means that the functional is almost determined by the lower bound for one-electron densities, and we would like to make use of this. We think that such a meta-GGA could be useful in its own right, but it could also be the basis for the construction of a more accurate fully non-local functional, for instance, via the Perdew–Zunger self-interaction correction,[47] which removes (on an orbital-by-orbital basis) the self-interaction error that all semi-local functionals have in their exchange part. A meta-GGA can be a component of a hybrid functional, where typically we mix in some exact exchange, and it can also be used in RPA to help make the short-range correction to that approximation.

Stretched bonds

What happens when we deviate from equilibrium bond lengths? If we have stretched bonds, the functionals typically need to be non-local.[40] When bonds are stretched, the exact exchange-correlation hole around an electron can spread out over two or more atomic centers. A classic example of this is stretched H_2^+. This is a molecule with two protons and one electron. The distance between the protons is greatly elongated, so that when the electron is sitting on one of the nuclear centers, half of its exchange hole is on that center and half is on the far-distant center. This situation cannot be described by a semi-local functional. It requires real non-locality, and without that non-locality, we have an incorrect description. The semi-local approximations localize the exchange-correlation hole around its electron, which means in this case the calculated exchange-correlation energy is too negative for stretched H_2^+. This is a typical self-interaction error, and it can only be corrected by a non-local approximation. These stretched bonds arise in binding energy curves. If we take a binding energy curve for a molecule and stretch out a bond length, and if an electron is shared between the two ends of the bond, there will eventually be a problem—a failure of the semi-local functionals.

Non-local stretched bonds also arise in the transition states of chemical reactions. When two molecules come together, they form a transition state, a complex or super-molecule. In this super-molecule, the bonds are typically weak and long. They are effectively stretched bonds, in which there is a significant non-locality. In order to describe the barrier height to a chemical reaction—the energy of the transition state, which is typically higher than the energy of the reactants or of the products—we must include real non-locality in our functional.

If we examine systems that have atomic-like *d* and *f* electrons, then even at equilibrium the *d* and *f* bonds can be stretched. Of course, the total bond is in equilibrium, but if we think of the bond as being made up of an *s* part and a *d* part, then the *d* part may actually be stretched, and so we also have such effects potentially occurring in these systems, which are often called "strongly correlated systems." The transition-metal oxides such as NiO and VO_2 are well-known examples. With the right amount of non-locality, we might also obtain a better description of strongly correlated materials, and this requires the two highest rungs of the ladder, either the hybrid functionals or the RPA-like functionals, to describe these stretched-bond, strongly correlated situations.

Hybrid functionals

Hybrid functionals[26–29] include an exact exchange component. One of the simplest of these is the PBE0 hybrid,[27–29] which mixes a fraction of exact exchange with a complementary fraction of PBE-GGA exchange and full PBE correlation. A typical fraction is 25% exact exchange. We can think of that

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

25% as a semi-empirical parameter, because the non-empirical argument for it[27] is not precise. This functional is usually much more accurate than PBE for the atomization energies of molecules and lattice constants of solids. It also improves the calculation of surface energies and energy barriers, compared to PBE. It gives improved but still inaccurate dissociation limits, but this is at least an improvement on semi-local functionals. We can also obtain a better prediction of energy gaps in materials,[48] because the potential in the hybrid functional is not a Kohn–Sham local potential but a mixture of a Hartree–Fock non-local potential with a Kohn–Sham local potential. (Hartree–Fock theory gives the energetically best non-interacting wave function to approximate the true correlated wave function.) The non-locality in the potential can mimic that of the self-energy and provide a reasonable description of bandgaps that would not be accurate even from the orbital energies of the exact Kohn–Sham density functional.

One criticism is that all successful hybrid functionals require at least one empirical parameter, and some have as many as 40. This means that we are missing some understanding of how to construct them. We are trying to develop non-empirical hybrids, for instance based on models for the exchange-correlation hole that use the exact exchange hole as an input.

Random phase approximation

The fifth-rung functionals are the RPA-like functionals.[16,30,31] RPA in a density functional context was proposed in the 1970s. It is a many-body approximation that includes unoccupied Kohn–Sham orbitals, unlike the first four rungs, which use only the occupied orbitals. RPA is nearly exact for the long-range part of the exchange-correlation hole, but the short-range part is in error. The total error from RPA is actually very large, –0.5 eV per electron in a typical system. That is a huge error, but almost all of that error cancels out from the energy differences we are interested in, if we make a change in the system that leaves the number of electrons unchanged. RPA can actually be useful for materials science, even without a correction, because of this remarkable cancellation of errors. For molecules, the error cancellation is somewhat imperfect,[30] so by itself the RPA is not very useful. However, Georg Kresse applied the RPA to solids in the past few years (e.g., Reference 30), and he found a much more nearly perfect error cancellation—similar to the error cancellation we had expected[49] to find in the 1990s. Kresse and others have found that RPA can yield very accurate lattice constants, surface energies, cohesive energies, and solid-state reaction energies. In our group, we have also found that RPA can yield accurate structures and accurate transition pressures[50] for structural phase transitions in solids. RPA also provides a good description of the long-range van der Waals interaction, which is missing from the first four rungs of the ladder. The RPA and the hybrid functionals—the fully non-local functionals—are available in VASP and other codes. We would like to construct a non-empirical non-local correction to RPA, which is required more for molecules than for solids.

Summary

Kohn–Sham density functional theory is currently the most widely used method for electronic structure calculation in materials science, condensed matter physics, and quantum chemistry, used not only by theorists but even by experimentalists. It is neither the only method nor the most accurate one, but it is often a very efficient and useful method. This method has a rigorous foundation for ground-state energies, ground-state electron densities, and geometric structures (i.e., equilibrium positions of the nuclei). It also has a hierarchy of approximations—the "Jacob's ladder"—of increasing sophistication for the exchange-correlation energy as a functional of the density. The ladder is still not complete. From my point of view, the first two rungs are complete, but we can still improve the third, fourth, and fifth rungs, and hopefully we can therefore make the functionals more accurate, more non-empirical, and more useful.

Acknowledgments

I would first like to thank the Materials Research Society for this Materials Theory Award, which was endowed by Toh-Ming Lu and Gwo-Ching Wang. I would like to thank the National Science Foundation for supporting my research over a long period. I also want to thank many collaborators; I cannot mention them all but I am especially grateful to Sy Vosko and David Langreth, who introduced me to DFT; they are no longer with us but not forgotten. I should also mention Alex Zunger, Mel Levy, Kieron Burke, Matthias Ernzerhof, Jianmin Tao, Gus Scuseria, and Adrienn Ruzsinszky. This work was supported, in part, by the National Science Foundation under Grant DMR-0854769 and, in part, by NSF EPSCoR and the Louisiana Board of Regents under the LA-SiGMA project (NSF Cooperative Agreement EPS-1003897).

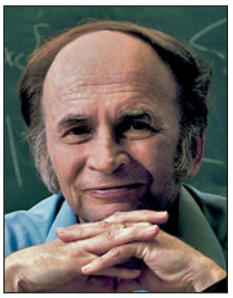
To view a video of John Perdew's presentation at the MRS 2012 Fall meeting, visit http://www.mrs.org/f12-mta-video.

References

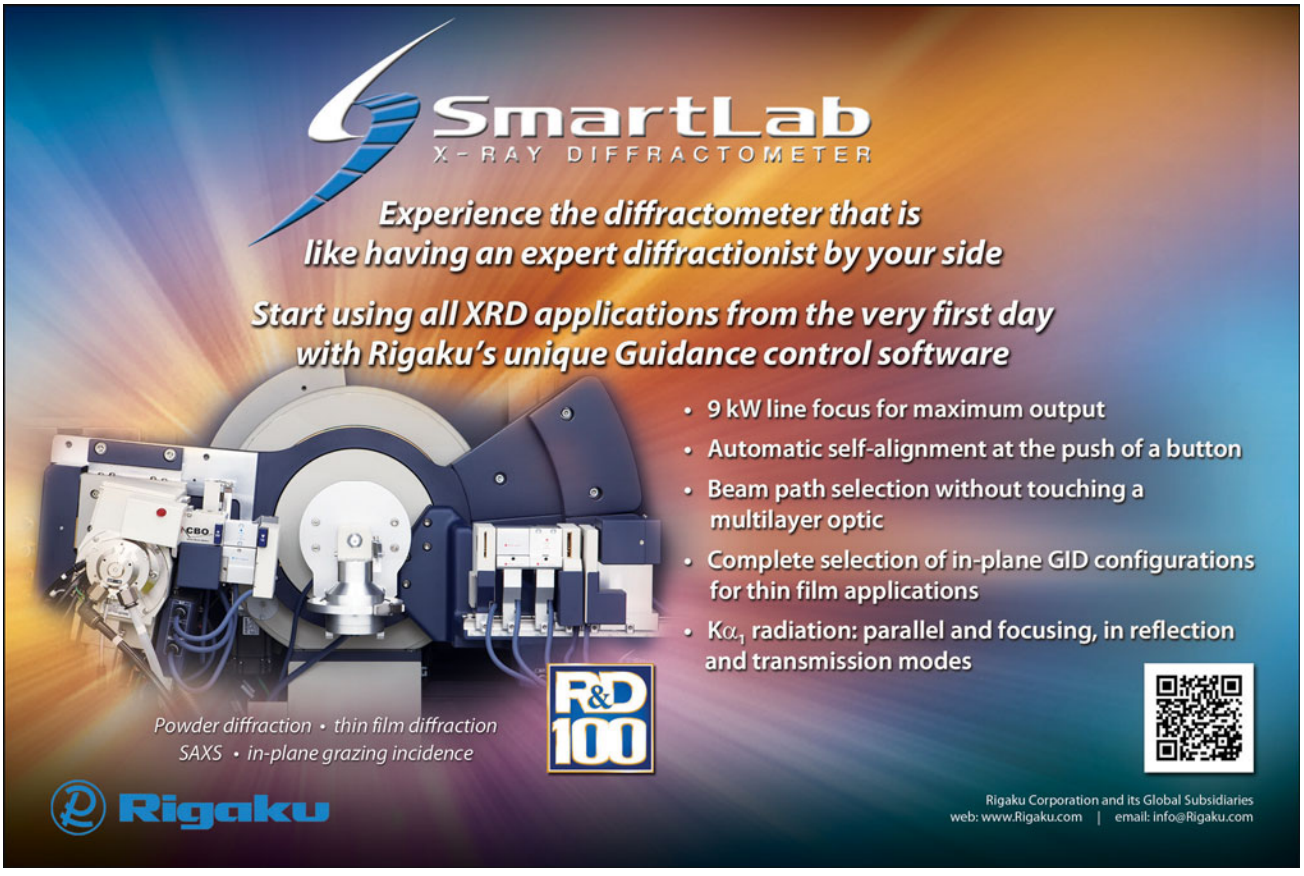
1. W. Kohn, L.J. Sham, *Phys. Rev.* **140**, A1133 (1965).
2. R.G. Parr, W. Yang, *Density Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989).
3. R.M. Dreizler, E.K.U. Gross, *Density Functional Theory* (Springer, Berlin, 1990).
4. C. Fiolhais, F. Nogueira, M. Marques, Eds., *A Primer in Density Functional Theory* (Springer, Berlin, 2003).
5. R.M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, UK, 2004).
6. E. Engel, R.M. Dreizler, *Density Functional Theory: An Advanced Course* (Springer, Berlin, 2011).
7. "18 Most Cited Physics Papers 1981–2010"; http://tulane.edu/sse/pep/news-and-events/upload/most-cited-papers-1981-2010.pdf.
8. W. Kohn, *Rev. Mod. Phys.* **71**, 1253 (1999).
9. S. Kurth, J.P. Perdew, *Int. J. Quantum Chem.* **75**, 889 (1999).
10. J.P. Perdew, Y. Wang, *Phys. Rev. B* **45**, 13244 (1992).
11. S.H. Vosko, L. Wilk, M. Nusair, *Can. J. Phys.* **58**, 1200 (1980).
12. D.C. Langreth, J.P. Perdew, *Solid State Commun.* **17**, 1425 (1975); *Phys. Rev. B* **15**, 2884 (1977).
13. O. Gunnarsson, B.I. Lundqvist, *Phys. Rev. B* **13**, 4274 (1976).
14. J.P. Perdew, K. Schmidt, in *Density Functional Theory and Its Applications to Materials*, V.E. Van Doren, K. Van Alsenoy, P. Geerlings, Eds. (American Institute of Physics, Melville, NY, 2001).
15. U. von Barth, L. Hedin, *J. Phys. C: Solid State Phys.* **5**, 1629 (1972).
16. D.C. Langreth, J.P. Perdew, *Solid State Commun.* **31**, 567 (1979); *Phys. Rev. B* **21**, 5469 (1980).

https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press

17. D.C. Langreth, M.J. Mehl, *Phys. Rev. B* **28**, 1809 (1983).
18. A.D. Becke, *Phys. Rev. A* **38**, 3098 (1988).
19. J.P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
20. A.D. Becke, *J. Chem. Phys.* **109**, 2092 (1998).
21. J.P. Perdew, S. Kurth, A. Zupan, P. Blaha, *Phys. Rev. Lett.* **82**, 2544 (1999).
22. J. Tao, J.P. Perdew, V.N. Staroverov, G.E. Scuseria, *Phys. Rev. Lett.* **91**, 146401 (2003).
23. J.P. Perdew, A. Ruzsinszky, G.I. Csonka, L.A. Constantin, J. Sun, *Phys. Rev. Lett.* **103**, 026403 (2009).
24. J. Sun, B. Xiao, A. Ruzsinszky, *J. Chem. Phys.* **137**, 051101 (2012).
25. J. Sun, R. Haunschild, B. Xiao, I.W. Bulik, G.E. Scuseria, J.P. Perdew, *J. Chem. Phys.* **138**, 044113 (2013).
26. A.D. Becke, *J. Chem. Phys.* **98**, 5648 (1993).
27. J.P. Perdew, M. Ernzerhof, K. Burke, *J. Chem. Phys.* **105**, 9982 (1996).
28. M. Ernzerhof, G.E. Scuseria, *J. Chem. Phys.* **110**, 5029 (1999).
29. C. Adamo, V. Barone, *J. Chem. Phys.* **110**, 6158 (1999).
30. F. Furche, *Phys. Rev. B* **64**, 195120 (2001).
31. J. Harl, L. Schimka, G. Kresse, *Phys. Rev. B* **81**, 115126 (2010).
32. S.-K. Ma, K.A. Brueckner, *Phys. Rev.* **165**, 18 (1968).
33. P.R. Antoniewicz, L. Kleinman, *Phys. Rev. B* **31**, 6779 (1985).
34. P.S. Svendsen, U. von Barth, *Phys. Rev. B* **54**, 17392 (1996).
35. M. Levy, J.P. Perdew, *Phys. Rev. A* **32**, 1010 (1985).
36. M. Levy, *Phys. Rev. A* **43**, 4637 (1991).
37. E.H. Lieb, S. Oxford, *Int. J. Quantum Chem.* **19**, 427 (1981).
38. O. Gunnarsson, B.I. Lundqvist, J.W. Wilkins, *Phys. Rev. B* **10**, 1319 (1974).
39. J. Jaramillo, G.E. Scuseria, M. Ernzerhof, *J. Chem. Phys.* **118**, 1068 (2003).
40. J.P. Perdew, V.N. Staroverov, J. Tao, G.E. Scuseria, *Phys. Rev. A* **78**, 052513 (2008).
41. S. Moroni, D.M. Ceperley, G. Senatore, *Phys. Rev. Lett.* **69**, 1837 (1992).
42. J.P. Perdew, A. Ruzsinszky, G.I. Csonka, L.A. Constantin, X. Zhou, O.A. Vydrov, G.E. Scuseria, K. Burke, *Phys. Rev. Lett.* **100**, 136406 (2008).
43. R. Armiento, A.E. Mattsson, *Phys. Rev. B* **62**, 10046 (2000).
44. Z. Wu, R.E. Cohen, *Phys. Rev.* **73**, 235116 (2006).
45. J. Sun, M. Marsman, A. Ruzsinszky, G. Kresse, J.P. Perdew, *Phys. Rev. B* **83**, 121410 (2011).
46. P. Hao, J. Sun, G.I. Csonka, P.H.T. Philipsen, J.P. Perdew, *Phys. Rev. B* **85**, 014111 (2012).
47. J.P. Perdew, A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
48. E.N. Brothers, A.F. Izmaylov, J.O. Normand, G.E. Scuseria, *J. Chem. Phys.* **129**, 011102 (2008).
49. Z. Yan, J.P. Perdew, S. Kurth, *Phys. Rev. B* **61**, 16430 (2000).
50. B. Xiao, J. Sun, A. Ruzsinszky, J. Feng, J.P. Perdew, *Phys. Rev. B* **86**, 094109 (2012). □



John P. Perdew is the Laura H. Carnell Professor of Physics and Chemistry in the Department of Physics at Temple University. He received a BS degree in physics and mathematics from Gettysburg College in 1965 and a PhD degree in physics from Cornell University in 1971. After postdoctoral work at the University of Toronto and Rutgers University, he joined the Department of Physics at Tulane University in 1977 and was promoted to full professor in 1982. At Tulane, Perdew served two terms as chair of physics and has supervised or guided the research of 15 doctoral students, 12 postdoctoral research associates, and two research assistant professors. He and his research group moved to Temple University in July 2013. He has published 265 research articles and presented 120 invited talks at conferences. In addition, Perdew is an elected Fellow of the American Physical Society and an elected member of the International Academy of Quantum Molecular Science and the National Academy of Sciences. He can be reached by email at perdew@temple.edu.



https://doi.org/10.1557/mrs.2013.178 Published online by Cambridge University Press