Review
Challenges for Theory and Computation

Karlheinz Schwarz

Institute for Materials Chemistry, Vienna University of Technology, Getreidemarkt 9/165,
A-1060 Vienna, Austria; kschwarz@theochem.tuwien.ac.at; Tel.: +43-1-58801-165301

Abstract: The routinely made assumptions for simulating solid materials are briefly summarized, since they need to be critically assessed when new aspects become important, such as excited states, finite temperature, time-dependence, etc. The significantly higher computer power combined with improved experimental data open new areas for interdisciplinary research, for which new ideas and concepts are needed.

Keywords: quantum mechanics; density functional theory; approximations; software; WIEN2k

1. Introduction

The role of computations for systems at the atomic scale (molecules or solids) has become routine for many standard applications, such as the interpretation of experimental data or providing a fundamental understanding of properties. Such calculations will also be needed in the future, and thus will remain important. However, there are new challenges, which we want to address with the focus on solids, interfaces, and surfaces, whose computational aspects were presented in a recent book chapter [1]. At present, it becomes important to scrutinize or reconsider the assumed basic assumptions and approximations used so far. They were needed to make computations feasible, but there are several cases that require new approaches, and thus one needs to go a step further. These cases include the atomic structure, quantum mechanics, the computational aspects (including software development), and the relation to experiments (and applications).

2. Presently Made Assumptions and Approximations

A series of assumptions and approximations are commonly made by simplifying and idealizing the complex solid-state materials so that we can simulate them and provide an understanding of their properties (e.g., [2]). We summarize these aspects in the next four subsections.

2.1. Atomic Structure

The early stages of solid-state computations were concerned with the study of relatively simple systems, such as a metal (e.g., Al) or an ionic solid (e.g., NaCl or CsCl), a semiconductor (e.g., Si), or a magnet (e.g., Fe or Co). For such systems, the assumption of a perfect crystal structure that can be presented by a unit cell was and is justified. This means that the unit cell—assuming periodic boundary conditions—is repeated to infinity in all three dimensions, and allows a representation in reciprocal space. These concepts are well known and made the computations of solids feasible. However, the interests have recently changed to more complex systems; for example, a crystal of finite size (nanocrystal), surfaces (e.g., in heterogeneous catalysis), impurities (doping), defects, or non-stoichiometry, as for example discussed in Ref. [1] (pp. 227–229). Additionally, the types of solids of interest have increased to cases such as superconductors, magnets, ferroelectrics, molecular systems, and all the way to proteins. This also means that a larger variety of chemical bonding now occurs, which includes metallic, covalent, ionic, but also van der Waals interactions.

2.2. Quantum Mechanics

The properties of crystals at the atomic scale are mainly determined by the electronic structure, requiring a quantum mechanical treatment. In the traditional approach, one goes from the time-dependent to the time-independent Schrödinger equation and neglects, as a first step, relativistic effects (i.e., we use Schrödinger's equation instead of Dirac's equation). Since the nuclei are much heavier than the electrons, one makes the Born–Oppenheimer approximation, in which the motion of the electrons is not coupled to the motion of the nuclei. This leads to Schrödinger's equation for the electrons, where the nuclei are at rest, taken at T = 0 K. For studying vibrations, we can perform several such calculations (with displaced atomic positions) and derive the dynamical matrix, from which phonons can be computed. There are also alternative ways based on perturbation theory. Usually the electron–phonon coupling is not included.

In the past, we had three different directions for a quantum mechanical treatment, with very little interactions between them. For molecular systems, the quantum chemists started with the Hartree–Fock (HF) method based on many-electron wave functions. In the HF scheme, exchange is treated exactly (by construction), but it does not include correlation effects due to an averaging over all other electrons for a selected electron. Correlation can be included in one of the post-HF schemes, like configuration interaction (CI) [3] or coupled cluster methods (CC) [4], both of which can reach almost exact results but with the drawback of a high computational cost. This limits the approach to rather small systems. In quantum chemistry, such schemes are called "ab initio" methods.

In the solid-state community, density functional theory (DFT) was the preferred scheme, in which electron density plays the key role. Walter Kohn received the Nobel Prize for the development of density functional theory, and thus it is appropriate to mention his significant contributions as summarized in an obituary for him [5], which also contains several relevant references. In contrast to the wave function-based methods, the fundamental idea of DFT is to replace the complete many-electron wave function with the much simpler ground-state electron density as the main variable. This is an enormous simplification because the density depends only on position (i.e., three variables). This opens the possibility of treating relatively large systems. However, the exact DFT functional is unknown, and thus approximations are needed—an active field of research. The simple DFT schemes are comparable to HF in terms of computer time, but only for localized basis sets (as used in chemistry) but not for plane waves (as used in solids). In contrast to HF, DFT methods include correlation, but treat exchange only approximately.

The third category is many-body physics (MBP) [6], which can treat complex phenomena such as highly correlated electrons, fluctuations or effects like electron–phonon coupling. Such schemes can describe complex situations, but are often based on parameters that are not directly derived for a real system.

Fortunately, the experts from these three categories for solving quantum mechanics (QM) problems started to collaborate, and they now benefit from each other. Each method has advantages and disadvantages. For example, one can include a fraction of HF exchange in DFT in so-called hybrid methods, or extract parameters from DFT for a many-body treatment. This combination of theory is one of the new challenges.

2.3. Computational Aspects and Software Development

For any calculation, one must as a first step define the atomic structure in an idealized form, with a unit cell or supercell that can present defects, impurities, interfaces, and surfaces (including vacuum). Periodic boundary conditions are assumed, which means, for example, that an impurity atom has a periodic image in the neighboring supercells representing an artificial order. The larger the supercell is chosen to be, the less is the effect of interactions between these artificial periodic images. These schemes make it clear that the atomic structure is always an idealization. The input data can come from experiment or are chosen on purpose for studying hypothetical cases. The advantage of theory is that the atomic structure—although idealized—is well-defined as input, in contrast to experiment, which can only approximately determine it.

The second step is choosing the quantum mechanical treatment as mentioned above, namely wave function based methods (HF and beyond) [3,4], DFT [1,2,7], or MBP [6]. In addition, the choice between all-electron schemes or valence electrons only (using pseudopotentials) must be made, but also specifying how relativistic effects (including spin-orbit coupling) shall be treated (as discussed, for example, in chapter 4 of Ref. [1] (pp. 234–238)). Then, a computer code is selected (see for example [2]) for solving the corresponding equations with a proper basis set (for example the WIEN2k program package that was developed in my group, see www.wien2k.at). The basis sets can consist of analytic (like Gaussian orbitals or plane waves) or numerical functions, or a combination of them. Depending on the property of interest, the convergence of a calculation needs to be tested; for example, in terms of the number of k-points (in the Brillouin zone) or basis sets (e.g., number of plane waves). For some properties (e.g., magnetic anisotropy energy), a high numerical precision is needed, since the energy difference may occur in the 10th decimal of the total energy.

It is very useful to have a large variety of computer codes in this field, since each of them puts the focus on different materials and properties. A certain code can be optimal for one case, but would not be a good choice for another. Sometimes we need a very high accuracy when we investigate fine details. In another case, a cruder calculation is sufficient for answering an open question. For the accuracy, we need a validation for well-chosen test cases. Reproducibility is important in this field. Different computer codes should give the same results, provided they use the same structure and (first principles) formalism (e.g., the same DFT functional) carried to full convergence. Recently, error estimates have been derived [7] for solid-state DFT calculations based on 40 different computer codes. These tests showed a very good agreement between the accurate codes (mostly all-electron full-potential codes), while deviations occurred for others (e.g., pseudo potential codes) which may be better in terms of efficiency. New concepts (methodology) can improve the efficiency, but new algorithms (e.g., for parallelization) can also be helpful, when properly chosen for the available hardware, ranging from laptops to supercomputers. With a more efficient computation, one can treat larger systems or explore more cases. The latter is needed, for example, in material design or optimization, where one can include all elements—irrespective of their abundance or environmental aspects—which are crucial for applications. Computations cannot find the optimal material, but they can "narrow design space" so that only those materials need to be synthesized and investigated which are predicted to have the desired property. For each open problem, we must find a good balance between accuracy and efficiency. The validation tests have shown that the deviation between accurate codes is often significantly smaller than the typical difference between theory (based on different functionals) and experimental data.

There are different ways of distributing a code: open source, access to the source code, or only executables; the latter is preferred by software companies. From a scientific perspective, the WIEN2k group favors making the source code available to the registered users. This policy has helped to generate a "WIEN2k community" of researchers (about 3000 groups around the world). Many of them have contributed to the development of the code in several aspects, such as bug fixes, adding or suggesting new features, and improving the documentation. Other developers also followed this strategy.

There is another problem that all computer code developers face—namely, user-friendliness. Based on the experience (from previous calculations) of the experts, one can provide many default options to make calculations easier—especially for novice users or experimentalists. However, there is also a drawback—namely the danger of using the code as a black box: "push a button and receive the result". Previously the users had to think about how to run the calculation and thus look at details instead of ignoring them.

2.4. Comparing Theory with Experiment

The importance of theoretical simulations has increased over the years due to the significantly improved computer power. Calculations can provide a basic understanding of structure–property relations (see for example [2]). In this context, it is often helpful to decompose the result into contributions for finding the driving force, which experiments cannot find. For example, one can

compute artificial cases like an impurity atom in a solid with and without relaxation of the atomic positions of the neighbors. From a comparison, we can clarify whether or not the relaxation is crucial for the studied property. The increased complexity of systems often makes computer graphic tools essential for analyzing many details of a computation. Take a unit cell with one thousand atoms and consider results as the band structure, density of states, or electron density. The data are stored somewhere, but the analysis needs new tools.

When deviations between a theoretical simulation and experimental data occur, we must critically scrutinize their origin.

- Did we properly model (idealize) the atomic structure (as discussed in Section 1)?
- Is the chosen quantum mechanical treatment, e.g., by DFT sufficient (Section 2)?
- Is full convergence reached in the calculation (as summarized in Section 3: in terms of **k**-points and basis sets)?
- Are there additional aspects which may affect the results, such as relativistic treatment, finite temperature, pressure, ground state, excited or metastable state?
- Are the assumptions or idealizations justified?

All five categories can cause an observed deviation between theory and experiment. Some aspects can be tested, for example by using a larger supercell, different DFT schemes, or repeating the calculation with more basis functions. With all the possibilities mentioned in the previous sections, it is often useful to combine different theories according to their advantages while keeping their disadvantages in mind. Even good agreement between theory and experiment may come from error compensation (too simple atomic structure but incomplete convergence). In some cases, parameters are used in a simulation, for example a Hubbard U in correlated systems, which may open a band gap. If one only adjusts U to fit the experimental band gap the agreement is obvious. However, when several experimental data agree for one chosen U, one gets at least a consistent picture which may be close to reality.

During the last decade, not only computation has improved significantly, but also the experimental techniques; for example, by improving the resolution (in space and time) or having better detectors. Recent developments (e.g., the use of short laser pulses) bring a new focus on the time dependence instead of time averaging assumed so far. We often need sophisticated experiments to ask the proper questions, which theory may be able to answer. Alternative explanations can be tested by simulations using predefined (partly artificial) test cases.

3. Discussion and New Challenges

Approximately 20 years ago, the fields of quantum chemistry, DFT, and many-body theory had hardly any cooperation between them, but this has fortunately changed. The strengths and weaknesses of the different approaches are recognized and mutually respected. One can solve complex problems only by close collaboration with the corresponding experts. Currently, computational chemistry and physics is often done within density functional theory, which occurs in a large variety of approximations. They need to be explored and validated by even more sophisticated methods, which, however, may be limited to relatively small system sizes due to their computational effort. The interdisciplinary nature of these issues is obvious, since the open challenges include chemistry, physics, mathematics, computer science, and materials science. By a combined effort of experts from all these fields, a substantial process has been achieved.

Another challenge comes from the experimental side. Take the work of Gatti and Macchi [8], who have presented a comprehensive overview of charge-density-related research that is closely related to DFT. For example, the multipolar model allows calculation of the static deformation electron density, and thus avoids the thermal smearing effects due to the atomic motions (Debye–Waller factor). Such presentations can be directly compared to DFT calculations that correspond to $T = 0$ K. New questions arise in connection with the time-resolved structural analysis based on the development of the X-ray free-electron laser making time-resolution of a few fs accessible [9]. This is a new challenge for theory, namely in terms of time dependence.

Let us illustrate what can already be done with a few examples. Time-independent DFT focuses on ground state properties. Formally speaking, we should not interpret the Kohn–Sham (KS) energies (in the form of the band structure) as excitation energies as discussed in [1,2]. However, this standard DFT single-particle model quite often describes excitations rather well. A proper DFT treatment for excited states is time-dependent DFT (TDDFT) [10]. In TDDFT one must make severe approximations by the choice of an approximate exchange-correlation-kernel, limiting the accuracy of this scheme. A properly chosen scheme (such as a core-hole calculation) allows the study of core-excitation spectra. The band gap is an important quantity for semiconductors or insulators, and can be calculated using an adjusted DFT functional (such as mBJ [11]). Recently, a careful discussion about the band gaps of solids was presented combining fundamental concepts (generalized Kohn–Sham theory) with applications for selected systems [12]. If DFT single particle theory is not sufficient, one can use the DFT orbitals as input for many-body perturbation theories, such as the GW approximation [13] for better quasiparticle energies, in which the self-energy $\sum$ is expanded in terms of the single particle Green's function G and the screened Coulomb interaction W. Another scheme is the Bethe–Salpeter equation (BSE) approach [14,15] to account for excitonic effects. Recently, the combination of TDDFT and excitons was presented in [16]. For highly correlated systems, which need a good description of the localized 3d or 4f states, inclusion of a Hubbard U into a generalized gradient approximation (GGA) may be sufficient for a proper description of the electronic structure (called GGA+U). However, sometimes one needs to go beyond this and use schemes like the dynamical mean field theory (DMFT) [17] to improve the agreement with experiments (mostly spectroscopy). Recently, even structure optimizations of correlated materials became possible by a combination of DFT (using WIEN2k) and the embedded DMFT [18]. Last but not least, the weak but sometimes important van der Waals interactions are usually not well described by standard DFT, requiring more sophisticated schemes [19]. In layered structures, DFT schemes (local density approximation-LDA, most GGAs, meta-GGAs, and hybrids as discussed in [1,2]) fail badly, but nonlocal versions demonstrate one way to proceed [20]. Such aspects become important in solid materials with a strong quasi-two dimensional regime [21].

It shall be mentioned here that there are other schemes to go beyond the topics present in this short review, such as molecular dynamics or thermodynamics, which are not mentioned here. Another direction is the ability to make use of extensive data (from theory or experiment) which are the key to the application of machine learning in materials science research [22]. The focus of this presentation is on solids, which are complicated systems, but they can be treated by a variety of methods. Many details are needed to make progress: concepts, a realistic atomic structure, accuracy, efficiency, validation of schemes, success or failure, trends, and predictions. In the past, idealizations and simplifying assumptions were necessary to make computations feasible. With the significantly improved computer power, it makes sense to critically analyze these assumptions, explore new concepts, and implement them into new (or existing) computer codes. There is no universal scheme that works for everything.

Conflicts of Interest: The author declares no conflict of interest.

References

1.　Schwarz, K.; Blaha, P. DFT calculations for real solids. In *Handbook of Solid State Chemistry*; Theoretical Description; Dronskowski, R., Kikkawa, S.H., Stein, A., Eds.; Wiley-VCH Verlag: Weinheim, Germany, 2017; Volume 5, pp. 227–259, ISBN 978-3-527-32587-0, doi:10.1002/9783527691036.hsscvol5022.

2.　Schwarz, K. Computation of material properties at the atomic scale. In *Selected Topics in Application of Quantum Mechanics*; Pahlavani, M.R., Ed.; InTech: Rijeka, Croatia, 2015; Chapter 10; pp. 275–310, ISBN 978-953-51-2126-8.

3.　Werner, H.-J.; Knowles, P.J. An efficient internally contracted multiconfiguration-reference configuration interaction method. *J. Chem. Phys.* **1988**, *89*, 5803–5814.

4. Bartlett, R.J.; Musial, M. Coupled-cluster theory in quantum chemistry. *Rev. Mod. Phys.* **2007**, *79*, 291.

5. Schwarz, K.; Sham, L.J.; Mattsson, A.E.; Scheffler, M. Obituary for Walter Kohn (1923–2016) *Computation* **2016**, *4*, 40.

6. Bloch, I.; Dalibard, J.; Zwrger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **2008**, *80*, 885.

7. Lejaeghere, K.; Bihlmayer, G.; Björkman, T.; Blaha, P.; Blügel, S.; Blum, V.; Caliste, D.; Castelli, I.E.; Clark, S.J.; Dal Corso, A.; et al. Reproducibility in density-functional theory calculations of solids. *Science* **2016**, *351*. doi:10.1126/science. aad3000.

8. Gatti, C.; Macchci, P. (Eds.) *Modern Charge-Density Analysis*; Springer: Berlin, Germany, 2000; ISBN 978-90-481-3836-4.

9. Schoenlein, R.W.; Chattopadhyay, S.; Chong, H.H.W.; Clover, T.E.; Heimann, P.A.; Shank, C.V.; Zholents, A.A.; Zolotorev, M.S. Generation of femtosecond pulses of synchrotron radiation. *Science* **2016**, *287*, 2237–2240.

10. Runge, E.; Gross, E.K.U. Density-functional theory for time-dependent systems. *Phys. Rev. Lett.* **1984**, *52*, 997.

11. Tran, F.; Blaha, P. Accurate band gaps of semiconductors and insulators with a semilocal exchange-correlation functional. *Phys. Rev. Lett.* **2009**, *102*, 226401.

12. Perdew, J.P.; Yang, W.; Burke, K.; Yang, Z.; Gross, E.K.U.; Scheffler, M.; Scuseria, G.E.; Henderson, T.M.; Zhang, I.Y.; Ruzsinszky, A.; et al. Understanding band gaps of solids in generalized Kohn–Sham theory. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 2801–2806.

13. Jiang, H.; Blaha, P. GW with linearized augmented plane waves extended by high-energy local orbitals. *Phys. Rev. B* **2016**, *93*, 115203.

14. Hetaba, W.; Blaha, P.; Tran, F.; Schattschneider, P. Calculating energy loss spectra of NiO: Advantages of the modified Becke–Johnson potential. *Phys. Rev. B* **2012**, *85*, 205108.

15. Laskowski, R.; Blaha, P. Understanding the $L_{2,3}$ X-ray absorption spectra of early transition $3d$ elements. *Phys. Rev. B* **2010**, *82*, 205104.

16. Turkowski, V.; Din, N.U.; Rahman, T.S. Time-dependent density-functional theory and excitons in bulk and two-dimensional semiconductors. *Computation* **2017**, *5*, 39.

17. Held, K. Electronic structure calculations using dynamical mean field theory. *Adv. Phys.* **2007**, *65*, 829–926.

18. Haule, K.; Pascut, G.L. Forces for structural optimizations on correlated materials within DFT+embedded DMFT functional approach. *Phys. Rev. B* **2016**, *94*, 195146.

19. Mori-Sánchez, P.; Cohen, A.J.; Yang, W. Many-electron self-interaction error in approximate density functionals. *J. Chem. Phys.* **2006**, *125*, 201102.

20. Rydberg, H.; Dion, M.; Jacobson, N.; Schröder, E.; Hyldgaard, P.; Simak, S.I.; Langreth, D.C.; Lundquist, B.I. Van der Waals density functional for layered structures. *Phys. Rev. Lett.* **2003**, *91*, 126402.

21. Thiel, S.; Hammerl, G.; Schmehl, A.; Schneider, C.W.; Mannhart, J. Tunable quasi-two-dimensional electron gases in oxide hertostructures. *Science* **2006**, *313*, 1942–1945.

22. Liu, Y.; Zhao, T.; Ju, W.; Shi, S. Materials discovery and design using machine learning. *J. Materiom.* **2017**, 3, 159–177.