Interface between quantum-mechanical-based approaches, experiments, and CALPHAD methodology


*Lawrence Livermore National Laboratory, Livermore, CA, USA
b Linköping University, Linköping, Sweden
c NIST, Gaithersburg, MD, USA
d SGF Scientific Consultancy, Aachen, Germany
e Royal Institute of Technology, Stockholm, Sweden
f 140 Clark Road, Brookline, MA, USA
g Max-Planck Institut für Metallforschung, Stuttgart, Germany
h TU Clausthal, Clausthal-Zellerfeld, Germany
i LTPCM-ENSEEG, St. Martin d’Hères, France
j University of Leeds, Leeds, UK
k Shanghai Institute of Ceramics, Shanghai, China

Received 14 February 2006; accepted 17 February 2006
Available online 22 March 2006

Abstract

The increased application of quantum-mechanical-based methodologies to the study of alloy stability has required a re-assessment of the field. The focus is mainly on inorganic materials in the solid state. In a first part, after a brief overview of the so-called ab initio methods with their approximations, constraints, and limitations, recommendations are made for a good usage of first-principles codes with a set of qualifiers. Examples are given to illustrate the power and the limitations of ab initio codes. However, despite the “success” of these methodologies, thermodynamics of complex multi-component alloys, as used in engineering applications, requires a more versatile approach presently afforded within CALPHAD. Hence, in a second part, the links that presently exist between ab initio methodologies, experiments, and the CALPHAD approach are examined with illustrations. Finally, the issues of dynamical instability and of the role of lattice vibrations that still constitute the subject of ample discussions within the CALPHAD community are revisited in the light of our current knowledge with a set of recommendations.

1. Introduction

Possibly the loftiest goal of computational materials science is the ability to design alloys with specific technological properties. To accomplish this goal it is necessary to have a thorough understanding of the fundamental mechanisms underlying materials behavior. In particular, one must understand the effects on alloy properties caused by changes in composition, temperature, or pressure (CTP), and how the addition of foreign elements affects alloy properties at the microscopic level. A great deal of information about the effects of CTP on alloy properties is contained in alloy phase diagrams which Massalski has called the “road maps” of the alloy designer [1]. In addition to equilibrium-phase information contained in phase diagrams, non-equilibrium processes are known to be crucial in determining materials properties. Transport, nucleation and growth mechanisms, martensitic transformations, metastable phases, and other phenomena...
affect both the microstructure of a system and its long-term phase and structural stability. A good understanding of these phenomena is indispensable in controlling the microstructure and the evolution of materials morphology that are known to play a significant role in the design of high-performance materials. Although there is little doubt about the necessity of acquiring knowledge of alloy properties at a fundamental level, there is a long way to go before we reach the age of application-oriented alloy engineering. At present, not only are there gaps in our understanding of the causes of equilibrium and non-equilibrium phenomena, but we have come to appreciate the interrelated nature of qualitatively distinct features such as thermodynamic and mechanical properties, as well as their connection to the underlying electronic structure of a material. Thus materials ductility, deformation mechanisms, creep properties, fracture toughness, tensile properties, and many other properties of a mechanical nature are often strongly correlated with particular alloy phases and chemical properties. It is known that these phenomena are strongly affected by the presence of specific forms of short- or long-range order in materials that, in turn, is intimately related to the underlying electronic structure. So-called environmental effects, such as the influence on mechanical properties of even small (ppm) additions of hydrogen, sulfur, or boron on the ductility of intermetallic compounds, are strongly coupled to the electronic structure of the system and can often be understood in terms of it. The modern study of alloy properties and the attempts to reach the ultimate goal of alloys-by-design have concentrated heavily on gaining a detailed knowledge of the electronic structure of a material and its effects on macroscopic and microscopic behaviors.

The study of the electronic structure of a material is invariably carried out within the formalism of quantum mechanics. Since the early 1930s, quantum mechanics has been leading to a deeper understanding of the behavior of electrons in atoms, molecules and solids. Attesting to the technological importance of alloys, the application of quantum mechanics to alloys dates almost as far back as the introduction of the discipline [2]. In carrying out theoretical studies of alloys based on quantum mechanics, one hopes to link the behavior of electrons – which in the case at hand can be taken to be the basic entity that holds matter together – to macroscopic behavior described by means of statistical mechanics and thermodynamics. Thus, calculations of the electronic structure, usually performed at zero temperature, provide information to complement thermodynamic data; the possible correlations between theoretical and experimental results can shed a brighter light on the underlying physics of alloys than would be obtained through theory or experiment alone. Now, the study of alloy thermodynamics and statistical mechanics is most often based on the so-called Ising model [3]. This model was introduced for the study of spin (magnetic) systems but has obvious formal similarities to alloys, although the basic physics of the two is very different. The modern theory of alloys is almost exclusively based on the coupling (or mapping) of the Hamiltonian describing the system at the electronic level to that associated with the Ising model. It is assumed that this mapping allows an application of the equilibrium thermodynamics and statistical mechanics developed in connection with the Ising model to the study of alloy phase stability and related properties. It is understood, of course, that this approach in its pristine form neglects the effects of kinetics in alloys that may provide mechanisms precluding processes that are determined to exist through the application of equilibrium thermodynamics alone.

Quantum-mechanical-based methods and their relations to the CALPHAD approach to alloy thermodynamic has been a recurring theme since practically the beginning of CALPHAD. Before focusing on specific aspects of the “interface between quantum mechanical-based approaches, experiments, and CALPHAD methodology”, let us briefly recall what the issues raised on this theme were during the four previous Ringberg workshops (RWs). In the first RW in 1995, on the topic of “Estimation of enthalpies for stable and metastable states” led by Juan Sanchez, it was recognized that large discrepancies between CALPHAD and ab initio structural energy differences existed. Already at that time it was concluded that “for an unstable phase the free energy itself cannot be properly defined” [4] (p. 485). Nevertheless, implementation of the CALPHAD approach to phase diagram computation requires that a method be devised to mathematically represent the thermodynamic functions of these unstable phases, although the quantities themselves may be void of physical meaning [4] (p. 485). Commonly practiced in the CALPHAD approach, the thermodynamic properties of non-equilibrium phases are inferred, whenever possible, by extrapolation of experimental data from a region in the phase diagram where the phase is stable. During the same workshop first-principles calculations were also found useful to calculate moments in the case of a magnetic system on the topic of “λ-transitions” led by Gerhard Inden [4] (p. 512). And on the topic of “Estimation of enthalpies and entropies of transition” led by Philip Spencer, both first-principles and tight-binding calculations were found useful to determine the formation energies for non-metallic materials [4] (pp. 544–545). During the second RW that took place in 1996, already a need for designing better physical models that describe the thermodynamics of end-members, and account for the electronic, strain energy, and chemical order effects in the case of substitutitional solid solutions was recommended [5] (p. 143). On the topic of “Thermodynamic modeling of selected topologically close-packed intermetallic compounds” led by Ibrahim Ansara, first-principles predictions of enthalpies of formation for simple structures (face centered cubic, fcc, body centered cubic, bcc, and hexagonal close-packed, hcp) and for complex structures such as the σ phase (e.g., of Fe–Cr) were presented [5] (pp. 182–183). During the same workshop, on the topic of “Order-disorder phase diagrams” led by Juan Sanchez, it was accepted that it was still a challenge for first-principles approaches to estimate, strictly speaking, the “other” components of the Gibbs free energy, such as vibrational, magnetic, and elastic contributions, although these terms may play an important role in the determination of the equilibrium phase boundaries [5] (p. 221). At that time, it was thought that a combination of
first-principles and CALPHAD approaches would yield significant advantages over the methods that were currently used. For example a combination of Cluster Variation Method (CVM) and CALPHAD was shown in the case of Ni–Al to definitely improve the description of phases that are related by an order–disorder transformation. Once again it was said that the “free energy \( G_f \) extrapolated into the unstable phase \( \phi \) of pure element I is to be considered just a convenient reference state without any particular physical meaning” [5] (p. 223). On the topic of “Solution thermodynamics of electronic materials” led by Alan Oates, it was noted that no better materials than those belonging to the class of electronic materials could be most appropriately amenable to first-principles electronic structure calculations. On this topic, and with some humor, a classification of methods was proposed that spans from first-principles to “first and a half”-principles (pseudo-potentials), to second-principles (semi-empirical tight-binding), to third-principles (empirical potential such as embedded atom method, EAM, potential that are fitted to experimental or first-principles derived properties), to finally fourth-principles (totally empirical potentials such as of the Lennard-Jones type) methods. It was also mentioned that molecular dynamics simulations usually rely on the results of third- and fourth-principles calculations, whereas Car–Parrinello simulations are based on first- and first-and-a-half-principles results [5] (p. 269). One of their conclusions was that a significant advantage of starting from sound physical formalism, rather than from ad hoc expressions to be fitted, is that it gives access to fundamental properties (other than energetics) that become increasingly relevant with the increase in sophistication of the software that is developed to predict alloy properties. In the third RW, held in 1997, although the emphasis was put on the applications of computational thermodynamics, the topic of “Models for composition dependence” led by Suzana G. Fries reemphasized the need for combining energetics from first principles and CVM for a proper description of chemical short-range order (SRO) in alloys [6] (pp. 33–37). Finally, during the fourth RW in 1999 a topic closer to what has been debated in the present RW was about “Using \( ab \) \( initio \) calculations in the CALPHAD environment”, led by Ben Burton [7]. Issues discussed in previous RWs were addressed with emphasis on insulating materials. Strategies were offered to estimate vibrational entropies and melting points from \( ab \) \( initio \) calculations, and to deal with SRO beyond pair correlations in CALPHAD calculations.

Because of the time constraint, only a few aspects that were thought to be of critical importance were touched upon during the 2005 RW. The paper is organized as follows. In Section 2, density functional theory that is the foundation of most electronic structure methodologies applied to condensed matter physics is briefly recalled. Due to the large variety of software that is now available to the non-expert for carrying out electronic structure calculations it was thought useful to assess their current underlying approximations, constraints, and limitations. In Section 3 we review the possible links that have been recently established between \( ab \) \( initio \), CALPHAD, and experimental results. In Section 4, despite the fact that the issues around dynamical instability and the role of lattice vibration in alloy stability have been settled in previous publications, the “lattice stability” issue has remained a “hot” topic in the CALPHAD community with little impact in physics and alloy theory. Therefore in this section the most salient findings on this theme are summarized with a set of recommendations for future work. In this section a brief update on the impact of vibrational effects on stability is also presented. The final conclusions are given in Section 5.

2. \( ab \) \( initio \) methods and codes: Overview, qualification, and examples

Technological progress requires materials with special properties. Accordingly, qualitatively new classes of materials have emerged with properties that are fascinating for scientists, as well as for engineers. Experimentally one can now alloy immiscible elements. It is possible to grow magnetic nanoparticles in diamagnetic or semiconductor matrices, as well as at surfaces, to prepare tiny nano-wires, and to synthesize crystals with forbidden symmetries, the so-called quasicrystals. These recent developments have led to increasing demands for a predictive power of the theory. Consequently, a new field of theoretical condensed matter physics has emerged, \( ab \) \( initio \) computer simulations of materials properties. The only required input into the simulations is atomic numbers of elements that build the material (though in principle the crystal structure can be determined in simulations, it is also quite often provided as input information to avoid unnecessary waste of computational resources). Simulations are based on the most fundamental laws of physics. Furthermore there is no need for any a priori experimental information, and there are no adjustable parameters in the theory. Hence, words like \( ab \) \( initio \) and first principles have found their place in the physics jargon.

Possibilities to study materials properties from first-principles electronic theory were enormously enhanced when density functional theory (DFT) and the local spin density approximation (LSDA) were formulated by Kohn and co-workers in the mid-60s [8,9]. In 1998 Walter Kohn was awarded the Nobel Prize for this theory. In the framework of the LSDA–DFT many practical problems of materials science could be solved successfully. Simultaneously, user-friendly computer codes for \( ab \) \( initio \) calculations were developed and released. Thus, electronic structure techniques are now employed for simulations not only by physicists, but also by chemists, geophysicists, biophysicists, metallurgists, and others.

At the same time, the use of different codes may, and often does lead to somewhat different results, and in certain cases the difference is dramatic. Perhaps the most well known example is the structural stability of pure Fe. Within the LSDA–DFT the theoretical predictions are qualitatively wrong, and bcc Fe is unstable. But the so-called generalized gradient approximation (GGA) [10] gives the qualitatively correct picture. In cases like this one can easily see that the latter approximation should be favored instead of the former. Unfortunately, this conclusion turns out not to be general. The GGA is known to improve
2.1. From the Schrödinger equation to the Kohn–Sham structure theory.

The general idea of ab initio simulations is illustrated in Fig. 1. Let us explain it more rigorously. The basis theory for ab initio simulations is quantum mechanics. To describe a system of atoms, one solves the Schrödinger equation (SE)

\[ -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + V(x_1, x_2, \ldots, x_{N_i}, R_1, R_2, \ldots, R_{N_i}) - \sum_I \frac{\hbar^2}{2M_I} \nabla_I^2 + \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|R_I - R_J|} \psi = E \psi \]  

(2.1)

where \(\psi \equiv \psi(x_1, x_2, \ldots, x_{N_i}, R_1, R_2, \ldots, R_{N_i})\); the indices \(i\) and \(I\) stand for number of electrons and nuclei, respectively; \(m_e\) and \(x \equiv r, \sigma\) correspond to electron mass, position \(r\), and spin \(\sigma\) respectively; \(M_I\) is the mass of the atomic nucleus at the coordinates \(R_I\), and \(Z_I\) is the charge of the nucleus. The potential energy operator \(V\) includes the electron–electron repulsion, as well as the external potential, e.g., electron–nucleus interaction.

It is generally believed that a complete solution of the SE for complex many-body systems such as a solid is not very useful, because the obtained information, in the form of the many-body wave function \(\psi\), is impossible to analyze [16]. In the context of the thermodynamic calculations this is probably not so obvious, as the most important information is contained in one number, the energy \(E\). But in any case, because the number of particles in any given system is of the order of \(10^{23}\), Eq. (2.1) cannot be solved exactly, and approximations must be introduced.

The easiest, and perhaps the most reliable, approximation present in all modern ab initio codes is the Born–Oppenheimer approximation, which allows one to study the electrons and the lattice dynamics separately. Because the electrons move much faster than the heavier nuclei, the later are assumed fixed in space while solving the SE for the electron subsystem. Within this approximation Eq. (2.1) is reduced to

\[ \left[ -\frac{\hbar^2}{2m_e} \sum_i \nabla_i^2 + V(x_1, x_2, \ldots, x_{N_i}, R_1, R_2, \ldots, R_{N_i}) \right] \psi_e = E_e \psi_e \]  

(2.2)

which is still a many-body problem. The density functional theory [8–10,16] reformulates it in terms of an effective one-electron problem. It is important to stress that DFT is a formally exact theory. Within DFT one solves the Kohn–Sham (KS) equations for a system of independent “electrons” (characterized by single-particle wave functions \(\phi_i\)), that have exactly the same density as the original system of real interacting electrons, i.e.,

\[ \left[ -\frac{\hbar^2}{2m_e} \nabla^2 + V_{KS}(x_1, R_1, R_2, \ldots, R_{N_i}) \right] \phi_i(x_1, R_1, R_2, \ldots, R_{N_i}) = \epsilon_i \phi_i(x_1, R_1, R_2, \ldots, R_{N_i}) \]  

(2.3)

The KS-DFT method is implemented in the majority of the codes for first-principles simulations of materials. It even made its way recently into quantum chemistry, where popular alternative approaches with acronyms such as Hartree–Fock (HF) and configuration-interaction (CI) have been used for some time to study complex molecules (N.B.: these methods would have limited-range applications in the case of solids).

Although the KS equations look very similar to the SE, they are much simpler to solve, because the effective potential \(V_{KS}\) depends on the coordinates of only one electron. According to the first Hohenberg–Kohn theorem it is a unique function of the electron density. Because neither the potential nor the density is known in advance, the KS equations are solved iteratively,
Fig. 1. (Color online) General idea behind ab initio simulations within density functional theory (DFT). When two atoms, A and B (left and right upper corners, respectively), form an alloy (middle panel), the electrons closest to the nuclei, the so-called core electrons (orbits inside dashed lines for corresponding atoms), remain atomic-like. But the outermost electrons, the so-called valence electrons, “leave” their respective atoms, forming bonds between atoms. According to DFT, the distribution of the electrons inside the alloy, the electron density $n(r)$, determines uniquely all the thermodynamic properties of the material. The task of ab initio simulations is to calculate $n(r)$ (middle plot of the bottom panel). From this one can determine the most important thermodynamic characteristics, like, for instance, the equation of states (left graph, bottom panel), mixing energy (right graph, bottom panel), etc.

starting with some reasonable guess for the charge density (often taken as a superposition of atomic charge densities). The second Hohenberg–Kohn theorem ensures that the converged solution corresponds to the charge density in the ground state. The total energy of the system is a unique functional of the density:

$$E_{\text{tot}}[n] = E_{\text{KS}}[n] + \int d^3r V_{\text{ext}}(r)n(r)$$

$$+ \int d^3r d^3r' \frac{n(r)n(r')}{|r-r'|} + E_{\text{II}} + E_{\text{xc}}[n].$$

(2.4)

Here $V_{\text{ext}}(r)$ is the external potential due to the nuclei and any other external fields, and $E_{\text{II}}$ is the interaction between the nuclei, that corresponds to the last term on the left-hand side of Eq. (2.1). Note that because this term corresponds to a rigid shift of the potential acting on the electrons, it can be easily accounted for within the Born–Oppenheimer approximation. We have also introduced the kinetic energy of independent electrons, $E_{\text{KS}}$, which is determined exactly within the Kohn–Sham method.

But of course there is a price to pay, namely, the exact relations between the effective potential and the total energy on one hand and the density on the other hand are not known. In particular, the exact form of the last term on the right-hand side of Eq. (2.4), the so-called exchange–correlation energy functional, is unknown. Basically, it incorporates all the difficulties, associated with the original many-body problem.

In the following we concentrate on a schematic description of three major “approximation circles”, see a schematic illustration in Fig. 2, that distinguish the different codes, and/or the most important options within a code. They consist in: (i) ways to approximate the effective potential in Eq. (2.3) and the exchange–correlation energy functional $E_{\text{xc}}[n]$ in Eq. (2.4), (ii) ways to expand the single-particle wave-functions while solving Eq. (2.3) numerically, and (iii) ways to handle the dependence of the effective potential on ions coordinates $R_j$.

2.2. Approximations for the effective one-electron potential: LDA, GGA, and beyond

The first approximation for the effective potential $V_{\text{KS}}$ in Eq. (2.3) was suggested in the original paper by Kohn and Sham [9], and is called the local density approximation (LDA).
Within the LDA the exchange correlation energy density at each point \( r \) in space for the real system with the charge density \( n(r) \) is assumed to be the same as that in the uniform electron gas with a constant charge density \( n_{\text{unif}} \equiv n(r) \). For the uniform electron gas the exchange–correlation energy can be calculated virtually exactly [10], and several useful parametrizations were suggested in the literature, see Ref. [10]. Though in general different parametrizations of the LDA functionals for the total energy and the one-electron potential lead to very similar results, one can strongly recommend to select one of the three most recent parametrized formulae proposed by Perdew and Zunger [17], Vosko et al. [18], and Perdew, Burke, and Ernzerhof (PBE) [19] because they are based on the many-body quantum Monte Carlo calculations of Ceperley and Alder [20]. Moreover, the PBE is consistent with one of the most well-known parametrizations for the GGA [19].

The local density approximation seems to be quite crude because real charge densities vary strongly, especially close to nuclei. Even its authors, Kohn and Sham, did not think that it was going to work in practice [9]. Still, numerous applications of the LDA were very successful, and lead to a wide acceptance of the DFT in the physics and materials science communities. However, one must be surprised by the success of the LDA rather than be disappointed by some of its (well- and less-well-known) failures, and the legitimate question here is “why does LDA work?” The question is still under debate in the physics community. One of the most reasonable explanations is the following. Let us consider one electron embedded in the sea of (other) electrons. Because of the electrostatic and quantum mechanical interactions, it repels other electrons, surrounding itself by a region with depleted electron density, the so-called exchange–correlation hole.

Though the LDA fails to describe the exchange–correlation hole in all its details, it describes its integrated average value exactly, and this was proven to be essential for the accurate predictions of materials properties. Still, in many applications the accuracy of LDA–DFT calculations is not satisfactory, and better approximations are clearly needed. Unfortunately, there is no systematic recipe for improving the approximations for the exchange–correlation energy and the one-electron potential. For instance, attempts to use a gradient expansion in terms of the charge density fail because a truncation of the expansion leads to unphysical results for the above mentioned exchange–correlation hole [21].

In the end, physicists succeeded in finding an approximation that accounts for the density gradients within the so-called generalized gradient approximation. In contrast to the LDA, there is no physical (model) system for which the GGA is exact (remember that the LDA is exact for the uniform electron gas). They obey certain known exact properties, but all known GGA functionals are basically constructed by hand, e.g. by cutting unphysical parts obtained by the gradient expansion [21]. Note that using different forms of the GGA may, and often does, lead to somewhat different results, say for the structural energy differences, mixing enthalpies, etc. Therefore, it is essential to specify in publications what parametrization of the GGA was used for the simulations.

The GGA does fix some remarkable problems encountered within LDA. For example, as was mentioned above, it predicts the correct ground state for iron. Generally, it gives bulk properties in better agreement with experiment, and therefore it can be safely recommended for the use in simulations of bulk properties. But we would like to re-iterate that one has to be careful with simulations of properties that involve open surfaces and volumes (surface energies, vacancies, etc.), where the GGA results often seem to be less reliable than the LDA ones. Once again, both LDA and GGA are approximations for the one-electron potential in Eq. (2.3). These approximations are known to be not suitable for calculating some properties such as band gaps in semiconductors, and for treating certain systems (the so-called correlated materials; in short, one has to be extremely cautious when performing DFT simulations for oxide materials and/or systems containing f-electrons). A lot of work has been done recently to design approximations that are suitable for treating these problems. One can mention the LDA + U scheme, the GW approximation, dynamical mean-field theory (DMFT), the exact exchange method, and the use of hybrid functionals [10]. Each of these methods has been proposed to solve a particular problem. Nowadays, substantial experience in solid-state physics is required for a proper selection and usage of these approximations.

2.3. Basis sets: Pseudo-potentials and all-electron methods

The main idea behind constructing different types of basis sets is schematically illustrated in Fig. 3. While most modern ab initio simulation packages allow one to carry out both LDA and GGA calculations, they differ in the way of numerically expanding the single-particle wave-functions \( \phi_i \) in Eq. (2.3). The background here is the following. Since a solution to the differential equations for complex systems, like real materials, is computationally impractical, one reformulates the problem in a set of linear algebra equations for which very efficient numerical methods currently exist (and are continuously improving). To do so one expands the one-electron wave-function \( \phi \) in Eq. (2.3) using a basis set \( \chi_j \):

\[
\phi(r) = \sum_j C_j \chi_j(r).
\]
regions: the potential spheres of radii $R_e$ have nontrivial solutions if $r \geq R_e$. It is easy to see that one arrives at a set of matrix equations that are most often plane waves. The whole space is divided into two regions: the core region (cf. regions inside the dashed lines for atoms A and B in Fig. 1), and inside these core regions real wave-functions are substituted by smooth pseudo-wave-functions in order to suppress the oscillations and to be able to truncate the plane wave expansion in Eq. (2.5) at a reasonable number of basis functions.

Then one substitutes Eq. (2.5) into Eq. (2.3), multiplies by the complex conjugate $\chi^*_j$, and integrates over the whole space. It is easy to see that one arrives at a set of matrix equations that have nontrivial solutions if

$$\det[H_{jk} - E O_{jk}] = 0. \quad (2.6)$$

Eq. (2.6), also called the secular equation, determines the energy eigenstates of the original Eq. (2.3). Note that the matrix elements in the secular equation, the Hamiltonian matrix $H_{jk}$ and the overlap matrix $O_{jk}$, are calculated straightforwardly in terms of the basis functions $\chi_j$. Strictly speaking, these matrices have infinite dimension. But in practice an expansion in Eq. (2.5) is truncated at a certain value $j_{\text{max}}$. The larger $j_{\text{max}}$ is, the more time consuming are the simulations. However, one must always converge calculations in terms of $j_{\text{max}}$ to obtain reliable results.

Note that the computational effort in solving the linear algebra problem scales as $M^3$ where $M$ is the size of the matrices $H_{jk}$ and $O_{jk}$. Therefore, one selects the basis functions as close to the solution of the Kohn–Sham equations as possible. The simplest basis set consists of plane waves, which are the solutions of the free-electron SE. In addition, in this basis set interatomic forces are calculated straightforwardly. Unfortunately, it is not easy to describe rapid variations of the charge density in the region close to the nucleus, the core region, in terms of a linear combination of plane waves. In order to circumvent the problem, one introduces so-called pseudo-potentials that have the same scattering properties for valence electrons, but artificially eliminate oscillations in the core region. Modern norm-conserving pseudo-potentials are very accurate, and they are used with great success in first-principles simulations [10]. Still, we believe that a verification of the pseudo-potential calculations by a comparison with all-electron calculations is highly desirable. At the same time, very accurate schemes that, to a large extent, merge the pseudo-potential idea with all-electron schemes is given within the projector augmented wave (PAW) method [22]. It is implemented in several simulation packages such as Vienna Ab-initio Simulation Package (VASP). It is strongly recommended to use PAW rather than the usual pseudo-potential technique if the former is available.

An alternative limiting situation as opposed to free electrons, is localized atomic-like orbitals. These are also used as basis functions in modern ab initio codes. Moreover, a very successful idea that allows one to create minimal basis sets is a complex basis function that is atomic-like close to the atoms, and free-electron-like between the atoms (in the interstitial region; see Fig. 3). Examples of codes where various basis sets and approximations are implemented are given in Appendix A.

2.4. Electronic structure calculations for ordered and disordered systems

So far we did not discuss the dependence of the effective potential $V_{\text{KS}}$ in Eq. (2.3) on ion coordinates $R_j$. The point here is that the cornerstone for the electronic structure theory of solids is the Bloch theorem. It is valid for systems with ideal three-dimensional periodicity, and allows one to reduce the Kohn–Sham equations to a problem for a single unit cell. As already pointed out, the computational effort in solving the linear algebra problem scales as $M^3$ where $M$ is the size of the matrices $H_{jk}$ and $O_{jk}$ in Eq. (2.6). In its own turn, $M$ most often depends linearly on the number of atoms in the unit cell $N$. Thus, conventional electronic structure codes scale as $N^3$. At the same time, almost all materials for technological applications have substantial deviations from this highly idealized model of a perfect periodic solid, and strictly speaking they have infinitely large “unit cells”.

The theory offers different treatments to the disorder problem. The methods can be classified into two major categories, the supercell-based methods and the methods based on perturbation theory. Let us consider the problem of substitutional disorder. In this case the two classes of methods differ in the implementation of the averaging procedure used to treat the disorder on an underlying crystal lattice, and this determines their advantages and limitations.

Within the supercell approach an alloy is viewed as a huge periodically repeated unit cell. Conventional band structure techniques are applied to the calculation of the electronic structure and the total energy of an alloy. Because of the $O(N^3)$-scaling limitation, one can currently treat supercells with several hundred atoms of different types. The question is now how one performs the necessary average over all possible configurations to determine the alloy properties of interest. The supercell method is based on the principle of spatial ergodicity, according to which all possible finite atomic arrangements are realized in a single infinite sample. What makes this principle work in practice is the fact that for a given physical property $P$, for instance, the total energy, all the correlations in the atomic distribution become unimportant at some distance and, hence,
the sample may be chosen finite. The problem is discussed in detail in Ref. [23].

The above statement has an important consequence. One must understand that the convergence of the calculations with respect to supercell size depends on the physical property considered. Because the distance at which correlations in the atomic distribution become unimportant is different for different properties. Thus, a supercell that gives converged results for the total energy may be too small to describe magnetic interactions, lattice vibrations, and optical and transport properties. Unfortunately, at present, the supercell size is most often chosen on the basis of available computer resources rather than on the basis of physical considerations. It is also important to underline that one can substantially improve the quality of supercell calculations by selecting a particular atomic distribution inside the supercell rather than doing a simple coin-flip construction. The approach is known as the so-called special quasi-random structures (SQS) method [24]. Because of the $O(N^3)$ scaling limitation, only a few SQS with $N \leq 32$ have been considered for binary alloys on fcc and bcc lattices, for two compositions, 50 and 25 at.%, and for a completely random atomic distribution inside the first few coordination shells. At the same time, a new generation of electronic structure methods, the so-called order-$N$ methods [25,26], allows one to calculate systems with up to 500 atoms even on a moderate workstation. With such a number of atoms in the supercell one may consider practically any random alloy composition with fixed correlation functions up to the sixth shell, including multi-component alloys [27]. Of course, values of the composition may only be rational fractions with an accuracy of $1/N$, but for many alloy problems such an accuracy is quite reasonable.

In contrast to the supercell approach, the methods based on perturbation theory perform the configurational average analytically. One of the most well-known approximations within alloy theory is the so-called coherent potential approximation (CPA) [28]. The CPA was originally introduced by Soven [29,30] for the electronic structure problem, and by Taylor [31] for phonons in random alloys. The success of the CPA is to a large extent associated with its formulation in the framework of multiple scattering theory developed by Györffy [32]. Combined with basis sets of the Korringa–Kohn–Rostoker (KKR) [28,32], Linear Muffin-Tin Orbital (LMTO) [33–35], or Exact Muffin-Tin Orbital (EMTO) [36,37] methods, the CPA has been used for calculations of bulk electronic structure, ground-state energetic properties, phase stability, magnetic properties, surface electronic structure, surface segregations, and many other characteristics of alloys.

The CPA is based on the assumption that the substitutional random alloy with, say, bcc or fcc crystal lattice, may be replaced by an effective medium with the same underlying lattice, the parameters of which must be determined self-consistently. A schematic representation of the basic idea behind the single-site mean-field CPA is illustrated in Fig. 4. In order to solve the electronic structure problem for the alloy, the effective medium is described by a (complex) coherent potential. In the single-site approximation it has to be determined self-consistently by the condition that the scattering of electrons of real atoms embedded in the effective medium vanish on average. Note that the CPA is a theory which is \textit{formally derived} as an expansion of the Green function for a disordered system [38–41], and it is possible to show that the corresponding series contains all diagrams describing electron scattering by a single site to infinite order.

The CPA equations can be numerically solved very efficiently (for the alloy problem, it is an $O(0)$ scaling problem!), and this turns out to be very useful in many applications. However, the CPA only gives a mean-field description of the electronic properties of the alloy. Moreover, because of the single-site approximation the original implementations of the CPA failed to determine the important Coulomb contribution to the total energy of an alloy coming from charge transfer effects, and they did not account for the effect of local lattice relaxations due to size mismatch between the alloy components. However, efficient and accurate schemes have been developed to account for charge fluctuations within the CPA [42–46], and the contribution to the total energy due to local lattice relaxations can be accounted for, for instance, within the effective tetrahedron method proposed recently by Ruban et al. [47].

Total energies calculated for completely random alloys by the supercell approach and by the CPA have been compared in several works, and very good agreement between the two techniques was found [48–51]. But because of the mean-field nature, the treatment of the short-range order effects within the CPA is impossible. This can be done using the recently proposed $O(N)$ methods, like the locally self-consistent multiple-scattering (LSMS) method [25] and the locally self-consistent Green function (LSGF) method [23,26]. In particular, the LSGF is an $O(N)$ method for calculating the electronic structure of systems with an arbitrary distribution of atoms of different kinds on an underlying crystal lattice. It is shown to be particularly suitable for the investigations of
random alloys that are modeled by large periodic supercells with several hundreds of atoms in the unit cell. The $O(N)$ scaling is achieved by associating each atom in the system with its so-called local interaction zone (LIZ) [25,52]. Inside each LIZ the multiple scattering problem is solved exactly. The accuracy of the LSGF calculations is controlled by the size of the LIZ, and its minimal size is ensured by embedding the LIZ in a self-consistent mean-field CPA-like effective medium.

Note that all the methods mentioned above are designed to calculate the total energy at zero temperature. To include temperature effects, one has to combine the electronic structure theory with methods of statistical physics.

2.5. Mapping of the electronic Hamiltonian onto an Ising model

Macroscopic order–disorder phenomena in alloys are usually described by an Ising-like Hamiltonian of the form,

$$
H = \sum_n \epsilon_n \sigma_n + \sum_{nm} \epsilon_{nm} \sigma_n \sigma_m + \sum_{nml} \epsilon_{nml} \sigma_n \sigma_m \sigma_l + \cdots \quad (2.7)
$$

where the “spin” variables, $\sigma_n$, vary over a specified domain and the $\epsilon_{i_1i_2\ldots i_n}$ describe the interactions (energies) associated with a cluster of sites $i_1, i_2, \ldots, i_n$. The $\epsilon_{nm}$, $\epsilon_{nml}$, etc. are commonly referred to as two-body, three-body, etc., interactions or potentials. The domain of $\sigma_n$ is defined by the problem at hand so that to describe a binary alloy $A_1-B_2$, in which atoms of types A and B are distributed over the sites of a lattice with corresponding probabilities $1-c$ and $c$, respectively, the $\sigma_n$ on site $n$ can take the two values $+1$ and $-1$, with probabilities $1-c$ and $c$. We note that the expression given in Eq. (2.7) is a generalization of that used by Ising to discuss the statistical mechanics of spins that could be oriented parallel or anti-parallel to a given direction (the $z$-axis), and which contained up to pairwise potentials. These general multi-site interactions can be determined from electronic structure calculations and can be used within the framework of equilibrium thermodynamics in connection with the Ising model to determine ordering tendencies as well as the $CTP$ phase diagrams of alloys. Once the interactions are calculated from an $ab$ initio methodology, the Ising model can be solved within a generalized mean-field approximation with the Cluster Variation Method (CVM) [53] or with Monte Carlo (MC) simulations [54]. In the CVM approach, at each temperature and alloy composition, the configurational part of the free energy is minimized with a Newton–Raphson technique with respect to the correlation functions defined as the following thermodynamic average product

$$
\xi_\alpha = \langle \sigma_1 \sigma_2 \cdots \sigma_\alpha \rangle \quad (2.8)
$$

where the maximum value of $\alpha$ correspond to the number of sites in the maximum cluster that characterizes the level of approximation of the CVM.

A number of procedures exist for extracting many-body interactions for alloys out of electronic structure calculations [55]. We briefly recall here those most commonly used in computational materials science. We classify these procedures into two main categories, (1) methods based on the electronic structure of random alloys and (2) methods based on the calculation of the electronic structure of specific ordered configurations of alloys. Methods such as the Generalized Perturbation Method (GPM) [56,57], the Embedded Cluster Method (ECM) [58,59], and Concentration-Functional Theory (CFT) [60] belong to the first category, while the so-called Connolly–Williams Method (CWM) [61] belongs to the second one. Of the ones listed in category (1), the first two correspond to perturbation-like expansions in direct configuration space, and the third in reciprocal space.

All methods in category (1) have been developed in connection with the Coherent-Potential Approximation (CPA) [28–30] for the study of the electronic structure of substitutionally disordered alloys, see Section 2.4. In this category, it is proved that the formation energy of an alloy $A_1-B_c$ in a given configuration $C$ can be written as

$$
\Delta E_{\text{Formation}}^C = E_{\text{Total}}^C - (1 - c^C) E_{\text{Total}}^A - c^C E_{\text{Total}}^B
$$

and the ordering energy associated with the configuration $C$ of the alloy, in the lowest order of perturbation, is

$$
\Delta E_{\text{Ord}}^C \approx \frac{1}{2} \sum_{n,m} (p_n^C - c^C)(p_m^C - c^C)V_{nm}(c^C) \quad (2.11)
$$

where $p_n^C$ is, at zero temperature, an occupation equal to 0 or 1 depending on the occupation of site $n$ by a B or A species. The total energy of the disordered state of the alloy $E_{\text{Total}}^c$ is approximated by the CPA medium, and is composition dependent. Since the expansion is derived with reference to the CPA medium, the interactions are composition dependent. In this category, the methods lead to uncoupling the total energy associated with a configuration $C$ in two contributions: the energy of the disordered state and the ordering energy, see Eq. (2.9). Since, within the GPM, the multi-site interactions and far-distant pair interactions can be easily calculated, there is no need to make any assumption on the truncation of the series in Eq. (2.7). Usually the expansion given in Eq. (2.11) converges rapidly with respect to the distance between sites $n$ and $m$ that characterizes the effective pair interactions $V_{nm}$, and the multi-site interactions beyond the two-body ones are negligible. However the convergence can and should be checked, since, in some instances (e.g., long-period superstructures), long-range interactions are required to explain the ordering trends. As said before, the interactions are composition dependent and have been used successfully to predict the ground-state properties of alloys. The scheme is represented in Fig. 5 (left panel).

It is worth mentioning a theorem that has been established in the context of a simpler electronic structure method based on the tight-binding (TB) approximation of the electronic structure [56]. The theorem says that the ordering energy of a (non-magnetic) transition metal alloy versus the number of valence d-electrons, $N_d$, exhibits two zeros besides the end
points, see Fig. 6. For example, one can simply conclude that Ni–Cu alloys should phase separate \((9 ≤ N_d ≤ 10)\), whereas Fe–Ti should display maximum ordering strength around equiatomic composition \((3 ≤ N_d ≤ 7)\).

It is implicit that for the methods belonging to category (1), the electronic structure method should rely on the Green function formalism, and therefore methods such KKR \([28,32]\) or TB-LMTO \([33]\) are most suitable. So far the limitations of the methods developed in this category are: (i) most codes rely on the Atomic Sphere Approximation (ASA) \([62]\) for the potential function, (ii) there are still pending questions on how to handle in a self-consistent and first-principles way charge transfer effects \([45]\), and effects of local lattice relaxations \([47]\) and (iii) most codes so far have been developed to handle simple structures such as fcc, bcc, and hcp.

In the second category of methodologies, it is assumed (ansatz) that the formation energy of an alloy \(A_1−cB_c\) in a given configuration \(C\) can be written as

\[
\Delta E^C_{\text{Formation}} \approx V_0 + \sum_{\alpha}^{\sigma_{\text{max}}} \sigma_1^C \sigma_2^C \cdots \sigma_\alpha^C V_\alpha
\]

where the spin variable \(\sigma_n^C\) is related to the occupation number \(p_n^C\) by \(\sigma_n^C = 2p_n^C - 1\). In this approach the mixing energy, i.e., the energy associated with the chemically disordered state of the alloy, is also assumed to be expanded in terms of the effective interactions \(V_\alpha\) according to

\[
\Delta E_{\text{Mix}}(c) = \Delta E_{\text{Dis}}(c) \approx V_0 + \sum_{\alpha}^{\sigma_{\text{max}}} [(2c - 1)^\alpha V_\alpha].
\]

Hence after selection of the maximum cluster(s) \(\sigma_{\text{max}}\), and of the ordered configurations \(C\) of the alloy, the interactions \(V_\alpha\) are determined by solving a set of linear equations once the formation energies of the alloy in each configuration \(C\) have been determined with an electronic structure code. The scheme is represented in Fig. 5 (right panel). Methods in this category (2) are usually wave based (such as pseudo-potential, APW, ASW, LASTO, LMTO) \([10]\) although, in principle, Green function methods can also be used. The advantages of the methods in this category are: ease of implementation, possible handling of complex structures, and accuracy afforded by full-potential methods. On the negative side, the selection a priori of the number and type of clusters, see the definition of \(\sigma\) in Eq. (2.12), and the selection of the ordered configurations that will be used to set up the system of linear equations (from which the interactions will be determined) are both ill-defined. Note that in this category the interactions are composition independent except via the volume effect, and usually the rate of convergence of the expansion given by Eq. (2.12) is rather small. In summary, although the differences between categories (1) and (2) seem rather subtle, they both lead to a completely different way of handling the statistical mechanical part of the problem. On one hand, in category (1), it is formally shown that the ordering-energy contribution to the total energy can be expanded in terms of pair and multi-site interactions, and the energy of the chemically random state of the alloy is directly given by the electronic structure method that can handle chemical disorder. On the other hand, in category (2), the total energy of the alloy is expanded in terms of multi-site interactions, and the energy of the disordered state is also described with the same interactions as those that describe order.

As a general statement, theories that are capable of giving more detailed explanations are automatically preferred \([63]\). Hence, it appears that the methods belonging to category (1) should be preferred since they are based on more solid formal constructs and truly possess predictive capabilities that have not been yet refuted by experiment. It has not yet been fully appreciated that the sole knowledge of the electronic structure properties of a chemically random state of an alloy carries
in it fundamental information on the ordering properties of the alloy; this relates to the fundamental power of statistical mechanics and the treatment of fluctuations (here of local alloy composition).

It should be noted that the methods in category (1) could be used as those in category (2) by assuming that Eqs. (2.9)–(2.11) hold, independently of the electronic structure method in use. In this case the set of energies \( \{E^{\text{CPA}}_{\text{total}}(c), V_{nm}(c)\} \) can be viewed as a set of parameters that have to be determined at each alloy composition) by generating an “appropriate” set of ordered configurations \( C \) (at each composition). Note that in the pair approximation the ordering energy is simply written as

\[
\Delta E^C_{\text{Ord}} \approx \frac{1}{2} \sum q^C_s V_s
\]  

(2.14)

where \( q^C_s = c(n^s_{BB} - cn_s)/2 \), with \( n^s_{BB} \) and \( n_s \) being the number of BB pairs and total number of pairs associated with the \( sth \)-neighbor shell. Hence, any ordered configuration \( C \) is simply specified by the set of geometrical factors \( \{q^C_s\} \), and the effective pair interactions (EPIs) \( \{V_s\} \) and the energy of the disordered state can be obtained by solving a set of linear equations, cf. Eqs. (2.9) and (2.14), once the formation energies have been numerically determined.

2.6. Recommendations

Because of the wide variety of ab initio software packages that are now available, it is important to know a priori the class of materials and the properties that will be targeted so that the right code with the proper approximations and the right capabilities can be selected (discussion with experts is highly recommended). In addition, during the selection process, it is important to inquire about code performance, system requirements, and support availability. Once a code has been acquired, it is important to:

(i) test it on systems such as Cu that have been reported in full details in the literature;
(ii) get acquainted with the code parameters (e.g., number of \( k \)-points, energy cut-off for plane-wave expansion, number of energy points for integration) and with code approximations (e.g., type of exchange–correlation potential, scalar relativistic or fully relativistic) and constraints such as spin polarization or not (magnetism collinear or not), core versus valence electrons;
(iii) test the validity of the code and of the parameter selection on properties (e.g., equilibrium properties, bulk modulus);
(iv) and most of all, check that the electronic structure itself makes sense (e.g., metal versus semiconductor, or insulator).

These are common sense suggestions that should be part of the qualification process through which one has to go when embarking in the “ab initio odyssey”. It should also be kept in mind that any ab initio code can deliver much more than numbers [64]. Among the properties that can be obtained with most electronic structure codes, it is worth mentioning:

- equilibrium properties, such as lattice parameter(s), structural energy difference, crystal-structure minimization;
- elastic properties, such as bulk modulus, elastic constants, phonon spectrum;
- electronic properties, such as density of states, charge density, and spectroscopic data;
- transport, such as DC and AC electronic conductivity and optical properties;
- magnetic properties, such as magnetic moments, susceptibility, and magnetic ordering;
- surface properties, such as structure, ordering, and energetics; and
- alloying effects on properties and, in particular, heat of formation (heat of mixing), heat of transformation, ordering energy, defect properties (e.g., stacking fault energy and anti-phase boundary energy), among others.

This richness of the acquired information can be used to verify the approximations and constraints that have been made during the calculations, and most importantly to provide a fundamental understanding of the origin of the properties that have been calculated. However as has been alluded to, despite the progress made in the field there are still challenging systems and problems that will require major formal development before meaningful predictions can be made. Among them, it is worth recalling the description of the liquid and amorphous states of matter, systems exhibiting large lattice distortions (such as some oxides and hydrides), some magnetic alloys (such as Ni–Cr) for which the magnetism is complex, and finally strongly correlated systems (such as some of the lanthanides, actinides, and oxides) for which there is still no fully ab initio approach that can properly treat electron correlations.

3. Links between ab initio, experiments and CALPHAD

In recent years, attempts have been made to make use of ab initio output, such as structural energies, heats of formation, heats of transformation, elastic and magnetic properties, in a CALPHAD description of alloy properties. This is the most immediate application of ab initio to CALPHAD. Furthermore, when the link between ab initio electronic structure calculations and statistical mechanics has been achieved, the information on thermodynamic properties as functions of alloy composition and temperature, and phase diagram results from CVM or Monte Carlo simulations are available. This can be included in a theoretical assessment “à la CALPHAD” of the thermodynamics of alloys (the same way an assessment is carried out within CALPHAD based on experimental data). This input from ab initio adds to the capabilities of CALPHAD to predict the thermodynamic properties of complex multi-component alloys. The situation of ab initio with respect to CALPHAD is schematically represented in Fig. 7. As indicated in this figure, the links between ab initio and CALPHAD provide fundamental information that is usually not accessible experimentally, and that can appreciably supplement the thermodynamic database. These links also establish a sound physical foundation for subsequent work on multi-component alloys, and for understanding the role played by solute in complex materials. Let us illustrate these links in the next two subsections.
The introduction of ab initio output, such as structural energies, heats of formation, heats of transformation, elastic and magnetic properties, can be used in a CALPHAD description of alloy properties. This is the simplest and most immediate application of ab initio to CALPHAD. This only requires to calculate the energetics (heats of formation) of existing or hypothetical compounds that can be directly input in a CALPHAD database. Additionally, in cases where a transformation occurs, heats of transformation can also be determined from ab initio without any difficulty. Let us illustrate this link with the case of Ni–Cr–Mo–W alloys. In CALPHAD databases no energetic information was available for the ordered Ni2Cr phase (C11 b or oP6) that is observed CALPHAD databases no energetic information was available. In cases when very little information is experimentally known on binary phase diagram, it was assumed that the Redlich–Kister parameters [70] of the two phases were the same except for an extra contribution to $0^{L_{\text{Ta,W}}} (or 0^{L_{\text{Mo,Ta}}}$, for the sake of simplicity. Hence, this procedure ensured that the $T_0$ location associated with the two phases (solid and liquid) at equiatomic composition was compatible with the experimental data on the liquidus–solidus lines [71]. Note that this latter assumption is not necessary, and, with different Redlich–Kister parameters for the two phases, a fit could have been performed with complete information on the assessed liquidus–solidus line in the whole range of alloy composition.

It was shown in both alloy cases that the overall fitting procedure leads to molar Gibbs energies and molar enthalpies of formation that reproduce fairly well the input ab initio information in a broad range of temperatures, and in the entire range of alloy compositions [67,68]. The error on the Gibbs energy between the CVM and CALPHAD results does not exceed 2% despite much larger errors on the enthalpy of mixing and the configurational entropy associated with differences between the CVM description of ordering and the single-site mean-field approximation afforded by the Bragg–Williams model. However the final CALPHAD phase diagram compares extremely well to the ab initio one as shown in Fig. 10. Note that the CALPHAD results of energetics and phase diagram agree overall very well with those obtained from ab initio at both low and high temperatures because the Redlich–Kister parameters have been made temperature dependent and the excess Gibbs energy is made composition dependent beyond the expression of the regular solution model, i.e., when higher-order Redlich–Kister coefficients $^PL$, with $p > 0$, are included. In other words, this description is equivalent to considering interactions that are temperature and composition dependent in a standard Ising model, and this versatility added to the single-site Bragg–Williams solution of the chemical order problem is the main reason for this agreement.

With this example, it is shown that ab initio results based on an electronic description of the alloy energetics and a generalized mean-field treatment of temperature effects can be cast in a CALPHAD formalism for subsequent prediction of the thermodynamic properties of higher-order multi-component alloys. Short-range order effects that indicate departure from ideal solid solution behavior are accounted for in an approximate way by assessing the ab initio results in the
Fig. 8. (Color online) Total energy (in kJ/mol) versus normalized lattice parameter (i.e., $a/a_{eq}$ where $a$ is the lattice parameter and $a_{eq}$ is the equilibrium lattice parameter) for Ni$_2$X (where X = Cr, Mo, W) of Ni$_2$Cr-type (C11b or oP6). The zero of energy is taken as the composition weighted average of the total energies of fcc Ni and bcc Cr (or Mo, W) (taken from Ref. [65]).

Fig. 9. (Color online) Calculated isothermal section of the Ni–Cr–Mo phase diagram at 893 K. Only the fcc matrix and the oP6-ordered phase are considered for this set of calculations (taken from Ref. [65]).

Fig. 10. Calculated Ta–W phase diagram obtained from the application of Thermo-Calc to the CALPHAD data (solid line), and compared with the CVM prediction (open circles), taken from Refs. [66,67].

same way as experimental phase diagram information, with a Redlich–Kister expansion that accounts for the temperature and composition dependence of the excess Gibbs energy. This conversion provides a robust and yet simple scheme that is consistent with standard assessment of multi-component alloy phase diagrams, as was illustrated in the case of Cr–Ta–W [67] and Mo–Ta–W [68]. It has been noted that the overestimation of a critical order–disorder temperature within the single-site Bragg–Williams approximation (as applied in a CALPHAD approach) is less severe for ternary alloys than for binaries. This should not come as a surprise since chemical order effects are becoming less important for higher-order component alloys than for binaries.

It can be concluded that the ab initio input to thermodynamic databases used in conjunction with the CALPHAD methodology can advantageously supplement the capabilities to predict the thermodynamic properties of complex alloys when experimental data is lacking. It is suggested that the links that have been established from ab initio to CALPHAD, see Fig. 7, can also be used in a reverse mode, i.e. from CALPHAD to ab initio, to test, validate, and challenge the approximations built in any first-principles approach to alloy stability and order.

As indicated in Fig. 7, we have shown that the links between ab initio and CALPHAD provide fundamental and valuable information that is usually not accessible experimentally, and that can appreciably supplement the thermodynamic database. Let us now examine how ab initio calculations can provide information for carrying out studies on the kinetics of phase transformations in unaries and alloy systems.
3.3. Ab initio contribution to kinetic database

The CALPHAD approach has been successfully extended and applied to the kinetic calculations for alloy systems. For example, in the software DICTRA [72] a quantitative analysis of phase transformation can be performed by combining the CALPHAD database with kinetics models such as the sharp interface model. Furthermore, recent progress in linking the CALPHAD approach with the Phase-Field Method [73, 74] enables us to describe and predict evolution processes of microstructural morphology during phase transformation in a multi-component system. The Phase-Field simulation combined with the CALPHAD database has been applied to a variety of phase transformations [75–78], and implemented in software package such as MICRESS [79]. In these kinetic calculations, one of the critical input parameters is the kinetic coefficient such as atomic mobility. The CALPHAD-type assessment for kinetic coefficients, which is suitable for the calculation of a multi-component system, has been suggested by Andersson and Ågren [80]. Within this methodology, the atomic mobility $M_A$ of an element $A$ is modeled according to

$$M_A = \frac{M_0^A}{RT} \exp \left( -\frac{Q_A}{RT} \right)$$  

(3.1)

where $M_0^A$ is the frequency factor and $Q_A$ is the activation enthalpy for diffusion. Both $M_0^A$ and $Q_A$ generally depend on composition, temperature and pressure. In the CALPHAD framework, these quantities are expressed by Redlich–Kister type polynomials [70] and are assessed based on experimental data such as tracer, intrinsic and chemical diffusion coefficients [81].

The number of experimental data for the kinetic coefficient is often scarce compared with the vast number of equilibrium data. It should be noted that the experimental work involved in obtaining kinetic coefficients with high accuracy has its challenges because of sample inhomogeneities and slow kinetics at low temperature. Since diffusion is a thermally activated process and diffusion coefficients depend exponentially on temperature, a small error in activation enthalpies results in a large uncertainty in the diffusion coefficients. Furthermore, the development of the kinetic database mentioned above requires information on atomic mobility in non-equilibrium structures that are difficult to experimentally measure. Ab initio calculation can help to reduce the number of free parameters in the assessment. Moreover, ab initio calculations can provide a clear physical picture of the microscopic diffusion mechanism. In the following, we briefly discuss ab initio works that are relevant for determining the kinetic coefficient, especially, activation enthalpy for diffusion in solids.

There are several possible diffusion mechanisms such as the direct exchange mechanism, ring mechanism and defect mechanisms via monovacancies, divacancies or self-interstitials. Among them, the monovacancy mechanism is dominant for most diffusion processes in unaries and alloy systems. When the diffusion is governed solely by the monovacancy mechanism with dilute vacancy concentration, the activation enthalpy consists of two parts, the formation enthalpy of the monovacancy, $H_f^V$, and the migration enthalpy of the monovacancy, $H_m^{\text{monov}}$. Electronic structure calculations based on the DFT have been successfully performed to study the formation energy of the monovacancy ($H_f^V$ at zero pressure). The adequacy of ab initio calculations to estimate for vacancy formation energy has been well demonstrated for simple metals [82–87]. Furthermore, DFT calculations based on full-potential methods proved their reliability in the case of transition metals [88–90]. Finally, DFT calculations have also been applied to the study of defect formation energy in intermetallic compounds and a detailed discussion can be found in Ref. [91].

These calculations are based on a supercell approach, i.e., the formation energy of the monovacancy is obtained from the difference between the energies for supercells with and without vacancy. The convergence of the energy with respect to the supercell size is, of course, important, and needs to be tested. Also, it is essential to take into account two types of relaxation effects on the vacancy formation energy, namely, volume relaxation and importantly, local structural relaxation around the vacancy [15,90,92,93]. In addition, it has been demonstrated that ab initio calculations of vacancy properties involve an error due to the existence of the internal surface, and correction for this error has been recently suggested [15,87]. Most importantly, it should be noted that the experimental value of $H_f^V$ is determined from high-temperature measurements, whereas the ab initio results generally represent values at 0 K. The importance of the temperature dependence of $H_f^V$ originating from anharmonic effects on lattice vibrations has been pointed out [87,94]. The temperature dependence of $H_f^V$ also stems from the electron excitation effect [93,95]. Hence, in principle, for a fair comparison between ab initio values and experimental data, it is necessary to take into account the anharmonic and electron excitation effects on formation energy.

Although the defect formation enthalpy is a static quantity, the migration enthalpy for diffusion is, in a strict sense, a dynamic one. The migration enthalpy in principle can be calculated from ab initio molecular dynamics simulations, which have been performed for self-diffusion in Si [96] and Na [97]. This method, however, carries with it a large computational burden to reach a reliable statistical accuracy. The most common way to obtain migration enthalpy is based on static calculations in the framework of Transition State Theory (TST) [98]. In the TST, the migration energy is expressed as the energy difference between two static states associated with the initial configuration and the saddle point configuration for migration process. Then, the key to a successful calculation of migration energy is the determination of the saddle point configuration for the microscopic migration process. For pure metals of simple crystal structure, the saddle point configuration can be often found on the basis of simple symmetry considerations. A more advanced method is the nudged-elastic band method [99]. The advantage of this method is that there is no restriction on the number of atoms involved in the migration process. Ab initio calculations of the migration
energy have been carried out in the case of, for example, vacancy-mediated self-diffusion in Al [100–102], Cu [103], Li [85,86], Na [86], Ta [95,104], W [93,104], and β-Hf [104], with acceptable reliability in the results.

Self-diffusion in pure metals is governed mostly by a single migration mechanism (at least at low temperature) and the experimentally observed migration energy is safely discussed in the light of the activation enthalpy for a single migration mechanism. However, the inter-diffusion process in alloys is significantly more complicated. The migration energy for a microscopic process crucially depends on the short- and long-range order as well as on alloy composition. The experimentally observed activation enthalpy consists of several contributions from various microscopic migration processes. In this regard, it is worthwhile to refer to recent work by Van der Ven and Ceder [105]. They presented an ab initio approach for calculating the vacancy-mediated inter-diffusion coefficient in binary alloys. In their work, the activation energies for several microscopic migration processes were firstly obtained from DFT calculations within the TST framework. Then, they described the configurational dependence of the activation energy by the “local” cluster expansion method [106]. In the local cluster expansion method, the activation energy (more precisely the “kinetically resolved activation barrier”) is described as a linear combination of polynomials of occupation variables, that are equivalent to the correlation functions in the conventional cluster expansion applied to the calculation of the configurational energy of perfect crystals. The configuration-dependent activation energy was employed in kinetic Monte Carlo simulations and the inter-diffusion coefficient was calculated within the Kubo–Green formalism [107]. By applying this method to Al–Li alloy system, they showed that the inter-diffusion coefficient is significantly affected by equilibrium short- and long-range order, and a simple diffusion model based on the single migration mechanism is no longer valid for the diffusion process in the alloys. This method should motivate further studies on ab initio alloy kinetics.

As mentioned above, a small error in the activation enthalpy results in a large uncertainty in the kinetic coefficients. A proper description of diffusion-controlled phase transformations requires a highly precise determination of activation enthalpy both experimentally and theoretically. The accuracy of ab initio calculations has been well demonstrated when compared with properly interpreted experimental data, and therefore ab initio studies should adequately supplement the kinetic database in the near future.

3.4. Recommendations

For all practical purposes, i.e. for the description of complex multi-component alloys, there is a need to have an accurate thermodynamic description of the binary subsystems. Until recently assessors mostly relied on input information from empirical methods such as the one proposed by Miedema in the early seventies [108] when experimental data were lacking. Nowadays the availability of sophisticated ab initio codes, see Appendix A, and the computational availability and power make the calculation of alloy energetics relatively easy and simple for an experienced user. Moreover, as was discussed in Section 2, compilations of energetic data and crystallographic information are now available either in the scientific literature [55,109–113] or, increasingly so, on various websites, cf. Appendix B.

Despite the limitations and constraints discussed in Section 2 as well as in the following section, it was shown that it is becoming possible to supplement a thermodynamic database with information obtained from ab initio calculations. Also in some instances when very little is known about an alloy system, ab initio predictions can be incorporated in the thermodynamic database and consequences on the properties of higher-order component alloys can be evaluated. In rare situations where predictions of ab initio binary phase diagrams are available, the results can be assessed in the CALPHAD context in the same way experimental phase diagrams are, with the advantage that during the evaluation process not only the phase diagram but also all thermodynamic functions are available for the assessment. Examples of both links were given. It is worth emphasizing that besides “numbers” ab initio methods do provide a fundamental insight into the role played by electrons in alloy phase stability and phase diagrams [64], and therefore an understanding of the trends, confirm or predict the existence of alloy phases, and give access to the topology of metastable phase diagrams. Finally, for composition–temperature–pressure phase diagrams that have been barely explored yet, ab initio calculations can provide valuable information on volume, elastic properties, and thermal expansion [114] that may be difficult to obtain experimentally. Similarly, for the study of layered nano-materials, ab initio results of surface and interfacial energies could be extremely useful.

It is worth mentioning that during the workshop, an “old experimental idea” from the 80s was put forward because of the direct connection with ab initio calculations [115,116]. By performing core-level shift calorimetry measurements with X-ray photo-electron spectroscopy (XPS) it is possible to relate the binding energies of deep lying core electrons of closed inner shells to cohesive energies [117] and heats of formation [115,116,118] in particular. Indeed these core electrons are sensitive to change in chemical bonding [119]. The shifts of these core-level binding energies as functions of alloy composition can be related to partial solution energies of alloys, and therefore to formation energies. This relation can be extended to ternary alloys for the partial solution energies. Based on the same idea, thermo-chemical data of metal surfaces and interfaces can be obtained from XPS experiments, although here, great care has to be taken to avoid surface contamination and other problems related to surface physics. However this latter information could find useful application in the study of small size systems. Hence, we recommend that this promising link be fully explored both experimentally and theoretically especially at a time where experimental data are lacking.

Because of time constraints not much has been said on the role that CALPHAD can play in ab initio studies. However, it is
important to allude to several links that will be worth exploring in the near future. First, the CALPHAD approach could be used to quickly survey systems that could be of interest for ab initio studies on the role of electron behavior on stability properties. Second, CALPHAD methodology could be used to supplement ab initio data with, for example, the vibrational contribution to the total Gibbs energy in the first-principles calculations of phase diagrams. And third, CALPHAD results could provide lower and upper bounds to various quantities that are presently calculated with ab initio methods, and therefore “validate” in some way the choices that have been made on the type of codes, the approximations, and the quality of the ab initio results.

Finally, despite recent attempts [105,120], effort will certainly be put in the near future into trying to augment our ab initio capabilities to provide information that enters the kinetic database, i.e., diffusion coefficients, mobilities and energy barriers, and this will allow us to make appropriate predictions about the kinetics of phase transformations with data generated from ab initio calculations.

4. Dynamical instability and lattice vibrations

4.1. How to reconcile ab initio and CALPHAD lattice stabilities?

Lattice stability (LS) is defined as the difference in Gibbs energies for a pure element based on two different crystalline structures. The definition has been extended to different chemical orders (compounds) in the case of multi-component systems. This concept was originally introduced by Kaufman [121,122] as a practical tool for assessing phase diagrams, and has found a formal significance in the introduction of the Compound Energy Formalism (CEF) by Hillert [123]. There is a consensus within the CALPHAD community that LSs that have been derived for most elements of the periodic table [124] were the results of extrapolation to end members of available thermodynamic data and phase diagram information. The introduction of these LSs was always considered as a matter of convenience, and their relevance was never an issue until comparison with ab initio data was made. Large discrepancies were observed for some elements among the transition metals, in particular those of columns VA and VIA (V, Nb, Ta, and Cr, Mo, W), and of columns VIlIA (Fe, Ru, and Os, and Co, Rh, Ir) [125]. During the first RW the notion of dynamical instability was introduced and already the validity of the ab initio data was questioned. Indeed once it was realized that some elements were displaying a dynamical instability, then it became clear that “for an unstable phase the free energy itself could not be properly defined” [4] (p. 485). Hence, dynamical instability prevents a meaningful use of the energetics calculated from ab initio calculations, and therefore even if agreement is found for some elements, it can only be fortuitous, unless the LS is associated with a metastable phase, and not an unstable one. A simple way to confirm this conclusion is to recognize that with most electronic structure codes only the electronic contribution to the “true” total energy is calculated. Within even the simplest Debye model for describing the vibrational contribution to the total energy, the zero-point contribution to the total energy per atom is given by (see, e.g., Refs. [126,127])

$$E_{zpm} = \frac{9}{8} k_B \Theta_D$$

where $k_B$ is Boltzmann’s constant and $\Theta_D$ is the Debye temperature that may be related to the bulk modulus, both being a measure of cohesion. For stable structures, the energy $E_{zpm}$ contributes very little to the total energy, and therefore can be safely ignored. The same would also apply to any phase (crystalline structure and atomic configuration) of an alloy, as long as it is a stable phase. However, severe complications arise if a phase is dynamically unstable under shear. The total energy still has a minimum for a certain volume, and a bulk modulus can be defined from the variation of the energy with volume under hydrostatic pressure. But it has no meaning to estimate a Debye temperature from that bulk modulus, and to use it in calculations of thermodynamic functions. This conclusion is not new and has been alluded to during the first RW, and subsequently in a series of papers [128–133]. However some of the obvious and immediate consequences have not been fully stated or appreciated. The first consequence is that the comparison between ab initio and CALPHAD LS data has no relevance for phases displaying a dynamical instability. The second consequence that has been largely ignored so far is that methods that rely explicitly on the energetics of unstable end members or ordered compounds to deduce energetic parameters for alloys at any composition are highly questionable. For example, in the Connolly–Williams approach (also referred to in recent years by the cluster expansion method) [61], chemical order in alloys is accounted for by expanding the total energy as a function of cluster interactions ($V_\sigma$), see Section 2.5. These interactions are obtained by inverting the set of equations that is established for a series of ordered configurations $C$ and the definition of a maximum cluster $\sigma_{max}$, and is given by

$$\Delta E_{conf}^C = \sum_{\sigma=0}^{\sigma_{max}} c_{\sigma}^C V_\sigma$$

where $c_{\sigma}^C = \sigma_1^C \sigma_2^C \cdots \sigma_\alpha^C$ is a multi-atom correlation function, see Eq. (2.12). However, if some of these configurations (including the end members) display a dynamical instability, the total energies associated with these configurations have no thermodynamically well-defined meaning, and therefore cannot be used to derive the cluster interactions (once again, only the electronic contribution to the “true” total energy is estimated). A final comment on this theme is worth making. The question “Why in some instances does it appear as though the comparison between ab initio and CALPHAD results is meaningful?” raises a more general one on finding a way to reach an agreement between ab initio and CALPHAD data on LS. Let us consider the case of Mo–Rh alloys for which the discrepancy in the ab initio and CALPHAD LS of Mo between the bcc ground state and the fcc phase is large [134]. Indeed, according to CALPHAD the energy difference between the two structures (extrapolated at zero temperature; in kJ/mol)
ranges from $-10.467 \ [121]$, $-28.00 \ [135]$, $-24.00 \ [136]$, to $-15.20 \ [124]$ whereas the ab initio LS (in kJ/mol) for Mo spans from $-41.25 \ (\text{LMTO--ASA+Combined Corrections})$, $-36.76 \ (\text{LMTO--ASA})$, $-62.97 \ (\text{fully relativistic TB-LMTO})$, $-33.66 \ (\text{LAPW})$, $-39.51 \ (\text{full-potential LASTO})$, to $-41.38 \ (\text{full-potential LPAW}) \ [134]$. Independently of the scattering energy, the fully relativistic version of the TB-LMTO within LDA and the SSOL solution database SSOL \ [137] is found between the CALPHAD LS (between the fcc and bcc phases) of $-17.00 \ [136]$, $-19.00 \ [124,135]$ and the ab initio value of $-16.28 \text{kJ/mol}$ (fully relativistic TB-LMTO) \ [134].

Fig. 11 shows the CALPHAD phase diagram with the hcp phase (that exists around equiatomic composition) “suspended” during the calculation, for the sake of clarity. The calculations have been performed with the Thermo-Calc software \ [69] and the SGTE solution database SSOL \ [137]. Note that the domain of stability of the bcc solid solution extends in a broader range of alloy composition than the fcc one. By making use of the fully relativistic version of the TB-LMTO within LDA and the coherent potential approximation (CPA) \ [33], the formation energy, $\Delta E_{\text{form}}^{\phi}$, of Mo$_x$Rh$_{1-x}$ based on one of the two lattices, $\phi = \text{fcc or bcc}$, is displayed as a function of $x$ in Fig. 12(a).

It is worth recalling that here the total energy has been calculated at each alloy composition, and therefore does not rely on the “legitimacy” of the values associated with the end members. According to what has been said above, one expects that above some alloy composition $x_1$ the fcc-based alloy is dynamically unstable whereas below a composition $x_2$ the bcc-based alloy is dynamically unstable. This means that $\Delta E_{\text{form}}^{\text{fcc}}$ for $x > x_1$, and $\Delta E_{\text{form}}^{\text{bcc}}$ for $x < x_2$ are quantitatively meaningful (provided one can ignore $E_{\text{zpm}}$). Let us consider for a moment how the LS values for Mo and Rh have been obtained within the CALPHAD approach. The evaluation methods rely mostly on extrapolation of the phase boundary lines, one of the oldest methods. It happens that not many phase diagrams display an fcc solid solution on the Rh side, or a bcc one on the Mo side. If we assume that the Mo–Rh phase diagram is the only one that has been considered in the evaluation protocol, the knowledge of (i) the energetics of each end member in its ground-state structure and the liquid phase (e.g. from experiment), (ii) some estimation of the entropy of melting for the phase other than the ground-state one, and (iii) the application of phase boundary extrapolations, allows one to estimate the LS for both Mo and Rh. It is apparent from the phase diagram that the bcc phase of Rh-rich alloys will be better extrapolated towards Rh than the fcc phase of Mo-rich alloys towards Mo. If one applies a similar extrapolation technique with, let us say, $x_1 \approx 0.4$ and $x_2 \approx 0.2$ to the ab initio results as shown in Fig. 12(b), then the LS values are about $+16.00$ kJ/mol for Rh and $+15.00$ kJ/mol for Mo; values that are in satisfactory agreement with the CALPHAD values. Note that in the case of Rh, since the value of $x_2$ is closer to 0 than $x_1$ is to 1 in the case of Mo), the extrapolation improves very little the LS value. However the agreement between the original ab initio value of LS and the assessed one for Rh still remains fortuitous (N.B.: once again, the energetics below the threshold $x_2$ has no physical meaning). This shows that

4.2. How do lattice instabilities affect CALPHAD calculations?

The Gibbs energy of a solid can be written as (see, e.g., Ref. \ [130])

$$G = H - TS = G_{\text{stat}} + G_{\text{ph}}$$

where $G_{\text{stat}}$ is the Gibbs energy associated with a static lattice and $G_{\text{ph}}$ is the contribution from lattice vibrations. Here we are interested only in ambient or low pressures so that the $PV$ term can be ignored in the Gibbs energy expression. Furthermore, we consider “high” temperatures $T$, i.e., $T > \Theta_D$ where $\Theta_D$ is a characteristic Debye temperature. For a single harmonic oscillator with frequency $\omega$, one has

$$G_{\text{ph}} = k_B T \ln[2 \sinh(h \omega/2 k_B T)] \approx -k_B T \ln(k_B T/h \omega).$$

For one mole of a solid, this contribution is given by

$$G_{\text{ph}} \approx -RT \ln(k_B T/h \omega_{\text{log}})$$

where $R$ is the gas constant and $\omega_{\text{log}}$ gives the logarithmic average of all phonon frequencies defined by

$$\ln(\omega_{\text{log}}) = \frac{1}{3} \int_0^{\omega_{\text{max}}} F(\omega) \ln(\omega) \, d\omega$$

where $F(\omega)$ is the phonon density of states.
A lattice is dynamically unstable unless $\omega^2(q, s) > 0$ for all phonons. Here $q$ is a wave vector and $s$ is a mode index (longitudinal or transverse, acoustic and optical), but the argument to follow holds even in the case of, e.g., an alloy where the frequencies of the atomic motions do not correspond to propagating plane waves but to more complex vibrational patterns. In a dynamically unstable lattice, the concept of vibrational entropy has no meaning. Despite the fact that the electronic contribution to the total energy can still be defined from ab initio electronic structure calculations, when the atoms are assumed fixed on a rigid lattice, the Gibbs energy $G$ is not a thermodynamically well-defined quantity. This is the cause of the discrepancy between CALPHAD lattice stabilities and ab initio cohesive energy differences, for instance as noted for tungsten based on bcc (stable) and fcc (dynamically unstable) lattices. This aspect has been discussed previously (see, e.g., Refs. [132,133]). In many cases an instability can be detected because an elastic shear mode is unstable. However there are also solids, which are dynamically stable under shear but have unstable phonon modes of short wavelength. The possibility of dynamical instabilities may seem to be a formidable complication in CALPHAD and ab initio studies, since in principle one has to check for every phase if it is dynamically stable. However, the existence of a dynamical instability, with precursor effects represented by an excess entropy has only a small influence on the phase diagram. This means that we can use the ab initio values for the enthalpy at intermediate alloy compositions, or even interpolate between the ab initio values for the lattice stabilities of the pure end members as shown in the previous subsection. The existence of an instability implies that there is another lattice structure that has a lower Gibbs energy. In practice the phase that eventually becomes unstable will not appear in the equilibrium phase diagram, when one is close to the instability. Therefore the excess entropy, which can be regarded as a precursor to the vibrational instability, varies slowly enough that it can be modeled by a standard CALPHAD approach. We remark that the modeling of an alloy that displays an instability can preferably be ended at a composition that is lower than the actual alloy composition where the instability arises.

4.3. Excess vibrational entropy in solid solutions

During roughly the last decade, there have been many studies of the effects of excess vibrational entropy, $S_{\text{Vib}}$, on the phase stability of intermetallics and alloys, both on the theoretical [139–151] and experimental [152–160] sides. A recent review, by van de Walle and Ceder [150], tabulates 19 systems that were modeled theoretically and 16 that were investigated experimentally.

The vast majority of these studies focused on the stabilities of intermetallic compounds, such as Ni$_3$Al, while studies of ionic insulating systems were limited to a single theoretical treatment of the system MgO–CaO [143]. A recent first-principles phase diagram (FPPD) calculation for the NaCl–KCl quasibinary system [161] indicates a very large $S_{\text{Vib}}$ effect: the calculated consolute temperature, $T_C$, for the miscibility gap is reduced by about 54% according to the formula:

$$\% \Delta T_C = 100 \frac{2(T_C - T_{C}^{\text{Vib}})}{(T_C + T_{C}^{\text{Vib}})}$$

where $T_{C}^{\text{Vib}}$ and $T_C$ are values for $T_C$ that were calculated with and without $S_{\text{Vib}}$, respectively. The theoretical results are displayed in Fig. 13 together with the experimental data of Refs. [162–165] for comparison.

This is in sharp contrast with previous studies of simple miscibility gap systems. Indeed, Tepesch et al. [143] reported a $\% \Delta T_C \approx 6\%$ for the isostructural MgO–CaO system; Silverman et al. [141,142] reported a $\% \Delta T_C \approx 3\%$ for GaP–InP; and for the wurtzite-structure systems AlN–GaN, GaN–InN and AlN–InN, Burton and van de Walle found $\% \Delta T_C \approx 6\%$, 12%, and 14%, respectively [166].

The ionic radius ratios, $\text{Na}^+/\text{K}^+ = 0.74$ and $\text{Mg}^{2+}/\text{Ca}^{2+} = 0.72$ [167]; are quite similar, so it comes as a surprise that $S_{\text{Vib}}$ is so much larger in NaCl–KCl than it is in MgO–CaO. Size mismatch, however, is only part of the story. Changes in
bond-stiffnesses, that occur in response to bond-length changes are also significant. Fig. 14 plots the stiffnesses of Na–Cl, K–Cl, Mg–O and Ca–O bonds as functions of composition $x_i$ ($i = \text{KCl or CaO}$). In the NaCl–KCl system, the Na–Cl bond stiffness goes from positive to negative at $x_{\text{KCl}} \approx 0.55$, whereas in MgO–CaO, the Mg–O stiffness becomes negative only at $x_{\text{CaO}} \approx 0.80$. This does not necessarily imply an unstable alloy, because the bending force constants (omitted in the graph) contribute to stabilize the solid. Nevertheless, a vanishing (or negative) force constant along the stretching direction indicates a relatively soft alloy with a large $S_{\text{Vib}}$. The large $S_{\text{Vib}}$ in NaCl–KCl does not occur in MgO–CaO because at $x_{\text{CaO}} > 0.80$ there are very few soft Mg–O bonds in the system, so the disordered Mg$_{1-x}$Ca$_x$O alloy is not significantly softened by Mg–O bonds.

In conclusion of this subsection, excess vibrational entropy is apparently a very important factor in the phase stability of the NaCl–KCl system. Including $S_{\text{Vib}}$ in the first-principles phase diagram calculation dramatically improves the agreement between theory and experiment, but does not produce quantitative agreement with respect to either $T_C$ or $x_C$. The improvement that is obtained by including $S_{\text{Vib}}$ only yields agreement with experiment that is similar to that achieved for the systems CaCO$_3$–MgCO$_3$ and CaCO$_3$–CdCO$_3$, without including $S_{\text{Vib}}$ [168]. Given the paucity of studies on ionic or structurally complex systems, it is premature to draw general conclusions about the benefits of including $S_{\text{Vib}}$ in any particular first-principles phase diagram calculation.

4.4. Recommendations

In the CALPHAD approach, the compound energy formalism requires that a thermodynamic description must be associated with any structure in the case of a unary system, and with any structure and ordered configuration in the case of a multi-component alloy system. In other words, it is assumed that the Gibbs free energy is “well behaved” even in the case of an unstable structure (or phase). Serious questions can be raised about “very hypothetical” (to say the least) phases that should be thermodynamically described within the CEF, such as Fe and Cr in the diamond structure or C in the $\sigma$ or bcc structure in the case of the Fe–Cr–C system [169] that are totally inconsistent with the present knowledge of bonding in solids. In addition, in the case of a dynamically unstable structure (or phase), the entropy $S$ is not defined, and therefore the Gibbs free energy is not either. This has been recognized since the first RW in 1995, and repeatedly afterward. On the rigorous side, for dynamically unstable phases, the ab initio total energy has no thermodynamically well-defined meaning, and therefore, on the practical side, for dynamically unstable phases, the zero-temperature energetics should not be used. One of the consequences of these findings is that methods that rely on the energetics of unstable end members or compounds are highly questionable. As was recalled in Section 4.2 if all the phonon frequencies are positive, then the system is stable (or possibly metastable) under small disturbances. The practical consequence of this assertion is that one should at least calculate the elastic constants (e.g., $C'$ in the case of cubic structures) if no knowledge has been acquired on the issue of dynamical instability. Furthermore, if the magnitude of a structural energy difference is “huge”, caution is advised. To the question: “Could a structure that is dynamically unstable at 0 K be stabilized at finite temperature?” the answer is obviously yes and examples include the case of Ti, Zr, and many other elements and alloys as well [170]. Under this circumstance, a possible treatment may be ab initio molecular dynamics simulations, but as of now, no result has been reported. To the question: “Could ab initio probe instability?” the answer is yes but it would require a tremendous effort that has not been put in yet.

On the practical side, as long as the CALPHAD information for a particular phase, obtained from a fit to phase diagrams, is not extrapolated and given physical significance in regions of alloy composition where that phase is dynamically unstable, it may be legitimate to use the traditional CALPHAD approach even in systems where some phases are not metastable but truly unstable. These phases will not show up in the equilibrium phase diagram, because they do not represent the lowest Gibbs energy. This was illustrated in realistic cases such as in Au–V alloys where similar phase diagrams could be obtained with very different lattice stability values for V [171]. Note that values of LSs are not so critical for binary alloys since what count are the formation energies. However for multi-component alloys, a spurious estimation of LS can give rise to “re-entrant” phases in multi-component phase diagrams [172].

Despite the absence of physical meaning to the ab initio energies of formation in regions where an alloy shows a dynamical instability, an extrapolation method similar to the one that CALPHAD is using can be applied to confirm or predict from ab initio the values of lattice stability as was shown in Section 4.1. In future applications this will require that the extrapolation protocol is specified in the same way a CALPHAD assessment is (or should be). It was shown in the case of Mo–Rh alloys that an extrapolation “à la CALPHAD”
of the ab initio data provides a good level of agreement between theoretical and assessed lattice stabilities.

5. Conclusions

During the 2005 Ringberg workshop, our task was to provide a current status of where ab initio methodologies were standing with their approximations, constraints, and limitations, and to examine the potential links that exist between ab initio, CALPHAD, and experiments. The issue of “lattice stability” has also been discussed, and although we believe the issue has been put to rest, its consequences have not been fully put in practice by both communities. Finally, we presented a revised version of how vibrational effects may impact stability properties especially in ionic systems. Each of the three sections that cover the topics mentioned above includes its own conclusions and recommendations.

Two final comments on future directions are in order. First, it is apparent that more and more ab-initio-derived information will flow into CALPHAD studies of statics and kinetics of phase transformations since experimental work in the alloy field is becoming sparse. Hence, it will become increasingly important to specify the origin of the information that is utilized to perform a CALPHAD treatment on alloys, if one wants to avoid another Tower of Babel. Probably one of the major restraints in absorbing the flow of ab initio data is the current inadequacy of the assessment techniques [172] in handling it efficiently and “objectively” so that new databases can be produced quickly and shared with others for further analysis and updates. Second, it has to be recognized that so far, besides ab initio “numbers” very little fundamental information has been passed on directly to CALPHAD modeling. For example, for electron phases such as the complex Hume-Rothery alloys [173], the introduction of relevant electronic parameters in the modeling would add to its robustness and its predictive capabilities, and therefore would create a knowledge-based approach to thermostatics and thermodynamics beyond phenomenology.

Among the subjects for which there was no time for discussions, it is worth mentioning in the context of the relation between ab initio methods and CALPHAD formalism, magnetic contribution to phase stability and phase stability at high pressure, and finally the need for formal improvements in ab initio methods that address, for example, electron correlations, temperature effects, and the treatment of the liquid state of matter, and their user-friendly implementation.

Acknowledgments

The authors are indebted to Matvei Zinkevich, Bo Sundman, and Fritz Aldinger for the organization of this workshop, to the German Research Foundation (DFG), the Scientific Group Thermodata Europe (SGTE), and the Max Planck Society (MPG) for the financial support, and to all the participants for stimulating discussions. The work of P.T. has been performed under the auspices of the US Department of Energy by the University of California Lawrence Livermore National Laboratory under contract No. W-7405-ENG-48. I.A.A. is grateful to the Swedish Research Council (VR) and the Swedish Foundation for Strategic Research (SSF) for financial support. G.G. acknowledges the support from the Swedish Foundation for Strategic Research under the project ATOMICS.

Appendix A. Ab initio software

Software packages are now available to perform electronic-structure-based calculations of materials properties. Most of these codes provide information about total energies, forces, and stresses, as well as calculate optimum geometries, band structures, optical spectra, and in some instances perform molecular dynamics. The following list only represents a sample of the codes that are available on the web.

A.1. Pseudo-potential-based codes

- VASP
  
  http://cms.mpi.univie.ac.at/vasp

  Pseudo-potentials and plane-wave basis set. VAMP/VASP is a package for performing ab initio quantum-mechanical molecular dynamics (MD). Ultrasoft Vanderbilt pseudopotentials (US-PP) or the projector augmented wave method (PAW) are used to describe the interaction between ions and electrons. Forces and stress can be easily calculated with VASP and it is used to relax atoms into their instantaneous ground state.
- CASTEP  
  http://www.tcm.phy.cam.ac.uk/castep  
  Pseudo-potentials and plane-wave basis set. This package can be used to calculate the properties of solids, interfaces, and surfaces for a wide range of materials classes including ceramics, semiconductors, and metals.

- ABINIT  
  http://www.abinit.org  
  Pseudo-potentials and plane-wave basis set. ABINIT can be used to find the total energy, charge density, and electronic structure of molecules and periodic solids.

- SIESTA  
  http://www.uam.es/departamentos/ciencias/fismaterial/siesta  
  Pseudo-potentials and LCAO basis set.

- Quantum-EXPRESSO  
  http://www.democritos.it/scientific.php  
  Pseudo-potentials and LCAO basis set. Open-source package for research in electronic structure, simulation, and optimization. The package contains three main codes: Pwscf (Plane-wave self-consistent field) for performing electronic structure calculations, structural optimization, molecular dynamics simulation, and for studying vibrational properties; FPMD and CP to perform Car–Parrinello variable-cell molecular dynamics simulations; and auxiliary codes (Pwgui, graphic interface and Atomic, a program for atomic calculations and generation of pseudo-potentials).

- FHI96MD  
  http://www.fhi-berlin.mpg.de/th/fhimd  
  Full-potential plane-wave code.

- Flair  
  http://www.uwm.edu/~weinert/flair.html  
  Full-Potential Linear Augmented Plane-Wave method for bulk materials and thin films.

- DFT++  
  http://dft.physics.cornell.edu/  
  Non-local potential plane-wave code within LDA, LSDA, or SIC. Plane-wave pseudo-potential and wavelet all-electron method.

- LmtART  
  http://www.mpi-stuttgart.mpg.de/andersen/LMTOMAN/lmtart.htm  
  Full-Potential Linear Response Linear Muffin-Tin Orbital (FP-LMTO) method.

- LMTElectrons (NMTASA, NMTCEL, and NMTPLW)  
  Full-Potential Linear Muffin-Tin Orbital method that comes in three versions: NMTASA (overlapping spheres), NMTCEL (unit cell divided in polyhedra), and NMTPLW (plane-wave code, non-overlapping muffin-tin spheres).

- LMTOPhonons  
  Full-Potential Linear Response Linear Muffin-Tin Orbital package designed to perform linear-response calculations of phonon spectra.

- LMTOMagnons  
  Full-Potential Linear Response Linear Muffin-Tin Orbital package designed to calculate dynamical spin susceptibilities of solids.

A.3. Others

- TB-LMTO–ASA  
  http://www.mpi-stuttgart.mpg.de/andersen/LMTODOC/LMTODOC.html  
  Tight-Binding Linear Muffin-Tin Orbital method within the Atomic Sphere Approximation (ASA), in the framework of the Green function formalism.

- TBMD  
  Tight-binding parameters and scalable molecular dynamics tight-binding code.

- SCTB  
  Charge self-consistent tight-binding total energy evaluation code.

- DENSEL  
  http://www.cmmp.ecl.ac.uk/~drb/Densel.html  
  $O(N)$ tight-binding code.
Appendix B. Databases

a. The following websites provide information on
   simple crystalline structures:
   http://www.intematix.com/
   Intematix Homepage
   http://alloy.cmu.edu/
   Alloy Database
   http://www.ca.sandia.gov/HiTempThermo/
   Thermochemistry Database for High-Temperature Materials Synthesis
   http://databases.fysik.dtu.dk/
   Materials Science Database

b. The following three websites provide information on
   crystalline structures. These are the crystal lattice structures
   from the Center for Computational Materials Science at
   NRL, the crystal data file (CDF) from NIST, and the
   Cambridge structural database (CSD):
   http://www.metallurgy.nist.gov/phase/solder/solder.html
   http://www.nist.gov/srd/
   Formation Energies for Binary Alloys
   NRL Code 6390
   http://www.physics.njit.edu/”mindlab/MaterialResearch/
   Databases/index_solids.html
   http://www-lab.imr.tohoku.ac.jp/”marcel/enthalpy/enthalpy.html

References

[4] B. Sundman, F. Aldinger, Ringberg workshop 1995 on unary data
   for elements and other end-members of solutions, CALPHAD 19 (4) (1995)
   433–571.
   the application of computational thermodynamics, CALPHAD 24 (1) (2000)
   15–94.
   514–525.
[23] I.A. Abrikosov, S.I. Simak, B. Johansson, A.V. Ruban, H.L. Skriver,
    (1990) 353.