Author: Teck Leong Tan

Phase Diagrams of Alloys from 1st-Principles Calculations

Abstract

This essay describes the formalism and methodology used for 1st-principles calculation of alloy phase diagrams via the cluster expansion method. A new methodology for the selection of optimal clusters in the expansion is described and its application on the Ni-V and Ti-Al binary alloys are presented.

1. Introduction and Background

The steady increase in computation power and improvement in algorithms has allowed scientists to probe the materials world via simulations. Although, experiments are indispensable in the study and characterization of materials, simulations are increasingly being used as a guide in the design of new materials, ahead of the real processes required to synthesize them. Due to their wide applications, metallic alloy is still a major field in materials research even though mankind had made use of them since ancient times. Alloys remain an interesting class of materials to study, since the properties of the alloy can be rather different from its parent metals. For example, the addition of small amounts of carbon (0.1%) to iron is enough to harden the otherwise soft iron significantly, due to the formation of the hard cementite phase, and the inclusion of chromium further gives rise to the "stainless" property of stainless steel. The formation of different phases with dissimilar properties is what makes alloying interesting and one is often interested in the atomic arrangements (structures) of the various meta-stable phases and the true ground state (stable phase) at different compositions. However, the experimental determination of the true ground state may be difficult, due to slow kinetics, i.e the system may take an unexpectedly long time to reach its equilibrium structure. This problem can however be circumvented in computer simulations, since we do not need to evolve the system in real time! Algorithms such as the Monte Carlo (MC) method, allows us to find the equilibrium structure rather efficiently. However, the correct prediction of the ground state configurations of an alloy depends very much upon the interaction energies being used. Traditionally, many computer simulations had relied on empirical data and parameters. For example, the semi-empirical Embedded Atom Method (EAM) potentials are still commonly used in the study of alloys and metals. Such potentials are often not "potable", i.e they are only good for a small class of materials system or the particular system of interest. As such, in recent years, much attention has been placed on "first-principles" methods which do away with empirical inputs.

The density functional theory (DFT) [1,2] provides the starting point for many 1st-principles calculations of materials systems including alloys. The approach is quantum-mechanical in nature and ideally, the end user only needs to provide the elements and their respective positions to extract the energy of the material system of interest. The DFT approach involves the solving of a set of self-consistent single particle equations with the effects of electron-electron exchange-correlation being approximated. Treatment of this topic is beyond the scope of this paper and the interested reader is directed to the excellent book by Martin R. [3]. Despite improvement in available DFT codes and computation speed, DFT calculations are still time-consuming and are often restricted to relatively small systems (< 200 atoms). Depending on the exact number of electrons treated and the desired accuracy, computation of energies for a mere 16-atom structure can take a day for a decent desktop! It is thus very inefficient to perform a ground state search for an alloy system by relying solely on DFT calculations, since the number of possible configurations of a 16-atom binary system is 2^{16}. Moreover, the thermodynamic properties and temperature dependence of the alloy system cannot be calculated by DFT and algorithms such as MC are required. The latter is a statistical mechanics method and can easily deal with hundreds or thousands of atoms. We thus need a multi-scale

approach to link the energies obtained from DFT calculations to the interactions required in MC simulations of alloy thermodynamics. The technique enabling such first-principle calculations of alloy thermodynamics is known as the cluster expansion, whose formalism is carefully laid out in [4, 5]. Even though the technique is straightforward, the construction of this expansion can be tedious and often relies on the researcher's intuition to guide the construction process. My paper shall focus on its use in the construction of alloy phase diagram [6, 7]. It should be noted that the temperature-composition phase diagrams produced are for the alloys in their solid state, since the MC is performed on a parent lattice (analogous to the Ising model done in our course). In addition, the order-disorder transition temperature of the alloy is the critical temperature of interest. Depending on the given interaction energies, atoms in a solid solution at a given composition may choose to form an ordered structure at low temperatures. The structure becomes disordered when the temperature goes beyond the critical temperature as illustrated by the schematic phase diagram of a binary alloy in figure 1. In essence, the 1st-principles simulations of alloy thermodynamics can be divided into 3 main parts:

(i)   A set of energies of ordered alloy structures are obtained from DFT calculations, $\{E_{DFT}\}$. Typically, the set contains 30 to 50 structures [6, 7].

(ii)   The Structure Inversion Method (SIM) (or otherwise known as the Collony-Williams [8] method) is used to convert the set of $\{E_{DFT}\}$ into a set of effective cluster interactions (ECIs). The (configuration) energy of the alloy system can then be written out as a cluster expansion, involving the ECIs and the cluster function.

(iii) The ECIs are used in an MC simulation which thermally equilibrates the alloy system to obtain thermodynamic properties of the system. The latter are used to construct phase diagrams.

As in any first-principles methodology, the most time consuming part lies in the quantum-mechanical DFT calculations and the above method has greatly reduced the number of DFT calculations required from 2^N to ~50 structures for a binary system with N (~1000) atoms. This makes the construction of alloy phase diagrams within a reasonable time period possible. In addition, it should also be noted the energies obtained from the DFT calculations can be stored in a database, should recalculations be needed later on. The rest of this paper will describe the cluster expansion method (focus is on step (ii)) prescribed in [6] and show its successful estimation of the order-disorder temperature for Ni_3V, as opposed to a previous cluster expansion. In addition, the paper will highlight the development of a computation package [7] aimed at producing alloy phase diagrams from another research group and their preliminary applications.
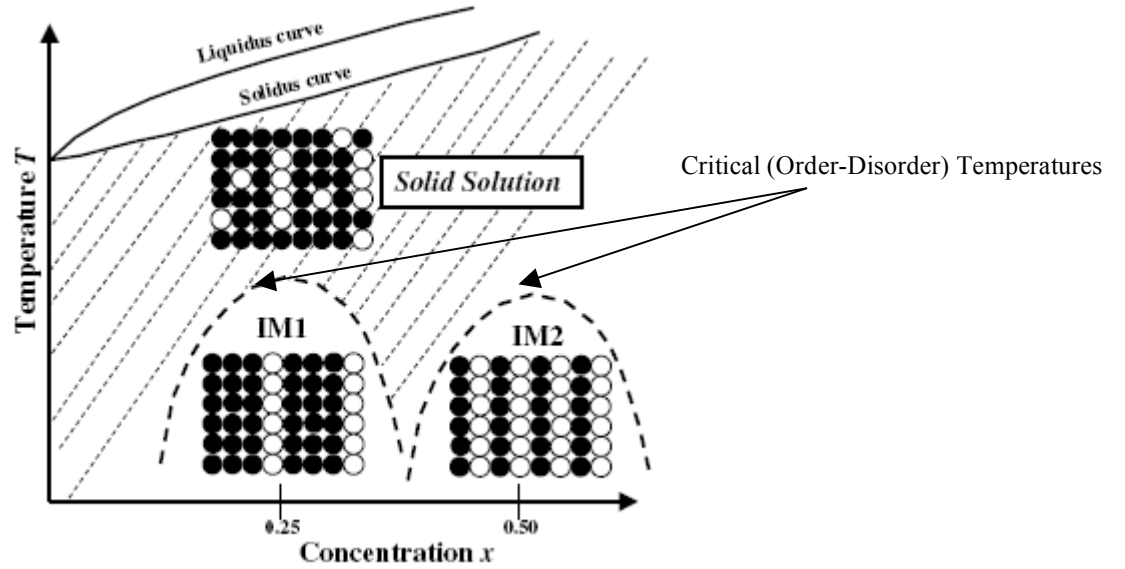
Figure 1: Schematic phase diagram of a binary alloy. Two ordered intermetallic compounds are found at x=0.25 and 0.5. The shaded portion represents the disordered alloy.

2. Cluster Expansion

The treatment of an alloy system can be made analogous to the Ising model which we learnt in class, with the spin variables being replaced by occupation variables and interactions are beyond that of the nearest neighbor. Given a parent lattice, the occupation at each lattice site can be represented by the occupation variable ξ^A_p which takes the value of 1 if the site p is occupied by an atom of type A and is zero if occupied by an atom of another type. The choice of occupation variables, ξ^A_p, over spin variables, s_p, (where s_p takes a value of +1 or -1 in a binary system) is that the former offers a more convenient treatment of the multi-component alloys. In such cases, the concentration of atom type A, c^A, is simply given by the thermal and site average the occupation variable, $<\xi^A_p>$. Further conversion are required if spin variables are used. In a binary system, we can write the occupation as simply ξ_p and the energy (per atom) of any atomic configuration σ can be written via cluster expansion as (see figure 2 for illustration):

$$E_{CE}(\sigma) = V^{(0)} + \sum_{n,f,d} V_f^{(n)} \Phi_{fd}^{(n)}(\sigma) \tag{1}$$

Where V and Φ denotes effective cluster interaction (ECI) and correlation function respectively. The summation is over clusters which are labeled by the index (n, f, d). The first 2 indices, (n,f), represents symmetry-distinct clusters, where n denotes the number of sites making up the cluster and f is an index which distinguishes symmetry-distinct clusters with the same number of sites. This system of labeling is illustrated in figure 3, on a face-centre cubic (FCC) lattice. The last index, d, simply distinguishes symmetry-equivalent clusters (i.e clusters with the same (n,f) but rotated with respect to one

another). For example, referring to figure 2, the 12 nearest neighbor pair will be labeled from (2,1,1) to (2,1,12) and the 6 second nearest neighbor pair will be labeled from (2,2,1) to (2,2,6). The n-site correlation function $\Phi^{(n)}_{fd}$ is obtained from the multiplication of the occupation variables of the sites:

$$\Phi^{(n)}_{fd} = <\xi_{p'}\xi_{p''}....\xi_{p^{(n)}}>$$
(2)

And the average is an ensemble average over the all clusters in the system with label (n,f,d). The correlation function thus takes on values between zero and one (inclusive). The cluster expansion can thus be viewed as an expansion with infinite terms, with Φ as the basis terms and V the corresponding coefficients [4,5]. If not truncated, the cluster expansion will be able to represent any $E_{CE}(\sigma)$ with an appropriate set of ECIs. In many cases, the series converges rapidly and the expansion can be truncated without serious lost in accuracy of the energy, $E_{CE}(\sigma)$. In many cases, only clusters that are relatively compact (e.g short range pairs or small triplets) are required and the truncated set of ECIs can be used in MC simulation, which involves the comparison of energies between different configurations. Now that we understand what exactly cluster expansion entails, two questions remains to be answered:

(i) How do we obtain the set of ECIs for our truncated cluster expansion series?
(ii) How do we decide which clusters should be included to yield an optimally truncated cluster expansion?
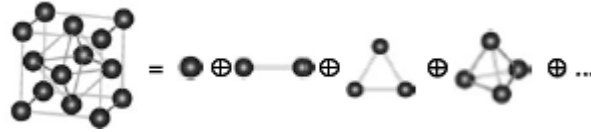

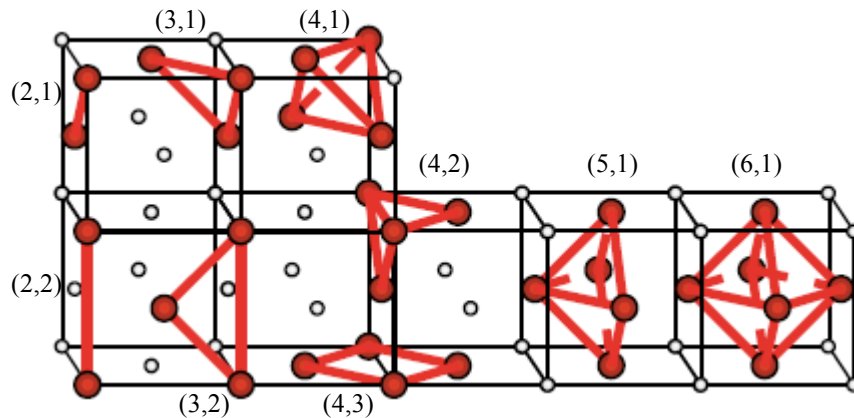Figure 2: Schematic illustrating the technique of cluster expansion.

Figure 3: Labeling of clusters for 2 to 6-body clusters for 1st and 2nd FCC neighbor, with (n,f) being the indices of the symmetry-distinct cluster. The 4-body upper set forms a locally complete basis in the range of 1st nearest neighbors, while the 6-body in both sets form the complete local basis in the range of 2nd nearest neighbors [6].

3. Structural Inversion Method (SIM)

The answer to the first question in the previous section is the Structural Inversion Method (SIM) [8]. Suppose we have a set of p structural energies from DFT calculations and q ECIs to fit too, we can write equation (1) in the following matrix form:

$$
\begin{pmatrix} E_{DFT}^{(1)} \\ E_{DFT}^{(2)} \\ \vdots \\ \vdots \\ E_{DFT}^{(p)} \end{pmatrix} = \begin{pmatrix} \Phi_{11} & \cdots & \Phi_{1q} \\ \Phi_{21} & \cdots & \Phi_{2q} \\ \vdots & \ddots & \vdots \\ \Phi_{p1} & \cdots & \Phi_{pq} \end{pmatrix} \begin{pmatrix} V_1 \\ \vdots \\ V_q \end{pmatrix}
\tag{3}
$$

In general, the number of data sets is more than the ECIs used for the fit, i.e $p \geq q$. One is thus faced with a "fitting" problem. The set of $\{V\}$ (ECIs) obtained should result in a set of cluster expansion energies, $\{E_{CE}\}$, which reproduces the original $\{E_{DFT}\}$ most closely, similar to a general least square minimization. I shall not dwell too much in this topic as it is a standard procedure in cluster expansion.

4. Selection of Optimal Cluster Expansion

The procedure for selection of the optimal set of clusters is what distinguishes the work in [6] from previous cluster expansions. During the construction of the cluster expansion, clusters are included sequentially and the following rules are adhered to during this process [6]:

(i)   If an n-body cluster is included, all n-body cluster of smaller spatial extent should be included too.
(ii)   If a cluster is included, all its subclusters should be included too.
(iii)   To prevent under fitting and over fitting, minimize cross-validation (CV) error which is given by:

$$
CV^2 = \frac{1}{p} \sum_{i=1}^{p} (E_{DFT}^{(i)} - E_{CE'}^{(i)})^2
\tag{4}
$$

where $E_{CE'}^{(i)}$ is the energy predicted by the set of ECIs fitted to p-1 DFT energies (with $E_{DFT}^{(i)}$ being excluded). Replacing $E_{CE'}^{(i)}$ with the cluster expansion energy, $E_{CE}^{(i)}$ in equation (4) yields the least square (LS) error. The latter is calculated by a fit to all p DFT energies. While the LS error measure the error in reproducing the energies obtained from DFT, CV error gives the uncertainty of predicted values. Hence, the CV error is a better measure of the "predictive power "of the ECIs obtained from structural inversion. In addition to rule (iii), rules (i) and (ii) ensures that important clusters of smaller spatial extent and/or smaller number of bodies are included first. Rule (i) is "physically meaningful" in the sense that large and extended clusters tend to have weaker interactions

than compact ones. For example, one would expect the magnitude of the interaction between the nearest neighbor pair to be stronger than a pair of atoms which are separated further away. The size of the cluster is defined by the longest pair in the cluster. Hence, referring to figure 3, clusters in the top row have the same spatial extent, likewise for the bottom row. Rule (ii) makes physical sense too. An n-body cluster can be interpreted as the interaction between an n'-body subcluster and an (n-n')-body subcluster. And it is not unreasonable to expect that the interaction within each of the subcluster is stronger than the coupling between the 2 of them [7]. Hence, the ECIs of subcluster are expected to be higher. Together, rules (i) and (ii) prevent the exclusion of more important clusters and transfer their weight to the less important ones. In addition, it leads to a simple algorithm to look through all physically sensible choices systematically [7].

5. Application on Ni_3V (Fixed Composition)

Using the set of energies from DFT calculations of 45 different Ni_3V structures, cluster expansion was performed and the optimal set of clusters consists of 3 pairs ad 3 triplets. The order-disorder temperature, T_c, turns out to be 1370 ± 180 K which agrees well with the experimental value of 1318 K. The error estimate used is the CV error. Figure 4 shows the trend of LS, CV scores and T_c as clusters are being added to the system systematically. As expected, the LS scores decreases monotonically as more clusters are added while the CV score shows a minimum point at 3 pairs and 3 triplets. This clearly shows that one should not use least square errors to determine the optimal cluster. If this were done, according the figure 4, the cluster will include up to the 7th triplet and the estimated T_c will be at 1800 K. The predicted value of 1370 K is a great improvement over previous cluster expansion calculations performed, which gave the T_c to be 1900 K [9]. The authors in [6] attribute the problem to the violation of both rule 1 and 3. The former gave rise to the inclusion of a longer-ranged 4-body cluster before the most compact one, while the use of least square fitting results in overfitting to the high energy $L1_2$ structure.
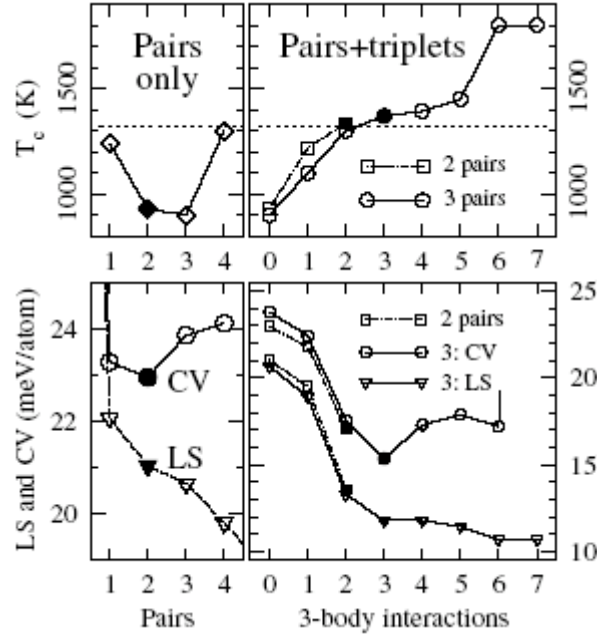
Figure 4: (Upper) The predicted and experimental (dotted line) T_c and (lower) the CV and LS error vs cluster truncation. Results are for CE with pairs only (left) and for 2 and 3 pairs with a number of 3-body ECIs (right). ECIs with minimal CV are given in solid symbols [6].

6. Application to Ti-Al System (Ti Rich Region)

Similar cluster expansion technique is also utilized by another research group in their development of a computational package to calculate phase diagrams [7,10]. The ATAT [10] package is aimed at simplifying and automating the procedure for doing 1st-principles thermodynamics via cluster expansion. The package adheres to the 3 rules specified in section 4. In addition, further checks are built in to ensure the prediction of the correct ground states and a scheme to improve the CV error through addition of structures to existing fit is included. The calculated phase diagram is shown vis-à-vis the experimental phase diagram in figure 5. The qualitative features of the Ti rich part of the phase diagram are reproduced and the ground states of TiAl and Ti_3Al are predicted correctly. However, the T_c calculated for Ti_3Al is 1850 K as opposed to the experimental value of 1450 K. The characteristics of the cluster expansion are given in Table 1. Using the weighted average of the CV scores, a CV error of 40 meV/atom is obtained. This corresponds to an error of ~480 K in the estimation of T_c. Hence, the experimental T_c does fall within the expected range of 1850±480 K.
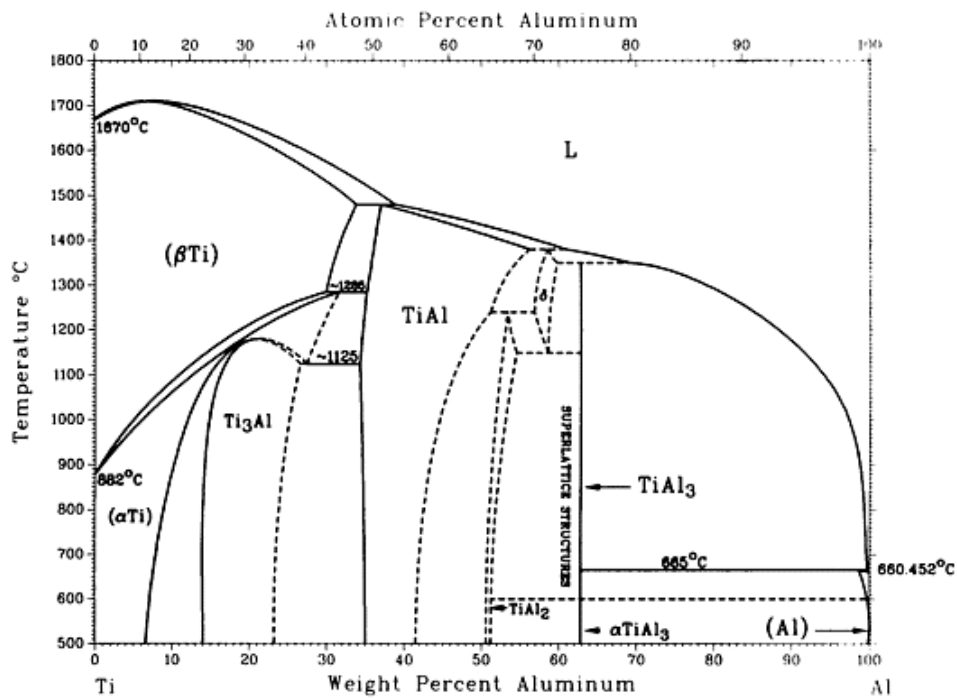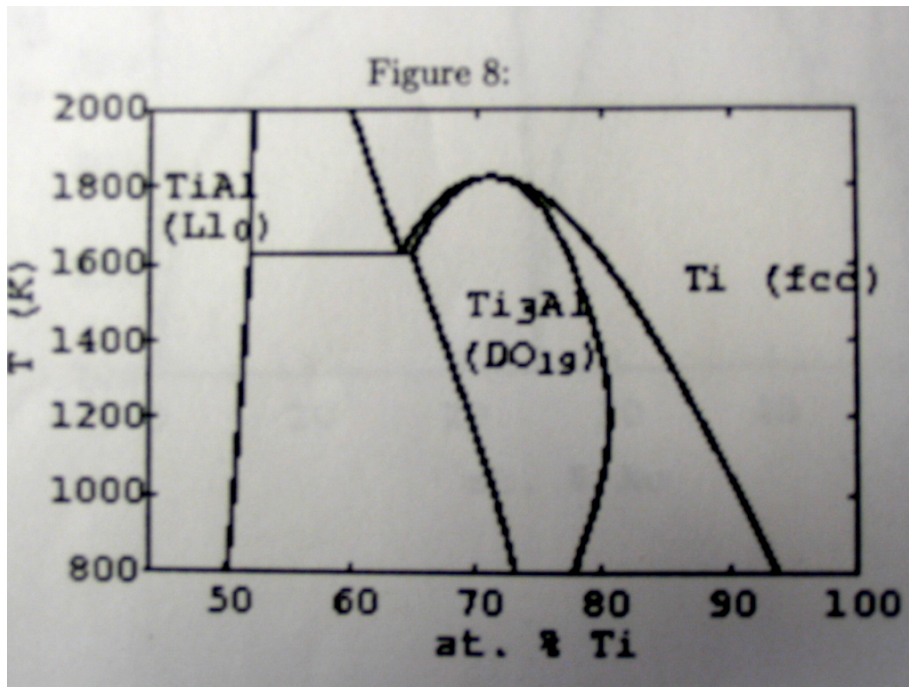
Figure 8:

Figure 5: (Upper) The calculated Ti-Al phase diagram (Ti-rich region). Concentration of Ti increases to the right. (Lower) the experiment phase diagram [11] with concentration of Ti increasing towards the left. Comparison is only valid in the Ti rich region.

Table 1: Characteristics of the Ti-Al Cluster Expansion

	Ti-Al (hcp)	Ti-Al (fcc)
No. of structures	55	23
CV score (meV/atom)	35	49

7. Discussion and Conclusion

In this paper, a methodology for obtaining reliable alloy thermodynamics via cluster expansion has been described. The study on Ni_3V in section 5 shows the importance of proper cluster selection. The authors in [6] were able to obtain a transition temperature close to that of the experimental value by adhering to the 3 physically sensible rules (spelt out in section 4) in their construction of the ECIs. A similar cluster expansion technique was also applied to the Ti-Al system by another group [7]. Although the calculated T_c of the Ti_3Al phase is not as close to the experimental value, the latter is within the error bars of the calculated T_c. Despite the fact that the calculated T_c are in reasonable agreement with experimental values, the former is somewhat higher in both cases. The origin of such an overestimation has not been unambiguously established [7]. One possible reason is the neglect of the effect of lattice vibration. The inclusion of lattice vibration should decrease the disordering temperature, since the atoms would possess more energy. Another possible candidate is the omission of weak but very long-range elastic interactions due to atomic size mismatch. Future first-principle calculations of alloy thermodynamics should address such issues.

References:

[1] Hohenberg P et. al., Phys. Rev. 136, 864B (1964).
[2] Kohn W et. al., Phys. Rev. 140, 1133A (1965).
[3] Martin R, "Electronic Structure: Basic Theory and Practical Methods", Cambridge University press 2004.
[4] Sanchez J M et. al., Physica (Amsterdam) 128A, 334 (1984).
[5] Sanchez J M, Phys. Rev. B 48, 14013 (1993).
[6] Zarkevich N A et. al., Phys.Rev. Lett 92, 255702 (2004).
[7] Van de Walle A et. al., J. Phase Equilibria 23(4), p348 (2002).
[8] Connolly J W et. al., Phys. Rev. B 27, 5169 (1983).
[9] Wolverton C et. al., Phys. Rev. B 49, 16058 (1994).
[10] Van de Walle A et. al., Calphad 26 (4), p539 2002).
[11] ASM handbook of alloys.