Non-statistical thermodynamic optimization

Advances in research related to the proposal and the previous Discovery Grant

It is difficult to say when exactly the CALPHAD (**CAL**culation of **PHA**se **D**iagrams) method was born and who the parents were. It is known, however, that the first issue of the CALPHAD journal appeared 30 years ago, in 1977. Since then, the ideas the method was based upon have been rapidly disseminating within the scientific and engineering community. Twenty one years later, in 1998, the thermodynamic, algorithmic and computational aspects of the method were summarized in the book "CALPHAD, calculation of phase diagrams: a comprehensive guide" [1]. This year, the book "Computational thermodynamics: the CALPAD method" [2] was published. Undoubtedly, the method has achieved an age of maturity, but does this mean that there is no need in its further development? For answering this question, let us recall what the previous advances were related to. 1. A development of new fast and robust algorithms to calculate equilibrium (including mapping), their program implementation and subsequent incorporation into existing software packages. 2. A development of new physically sound thermodynamic models capable of describing the properties of intricate phases (*e.g.* phases with multiple sublattices, liquids prone to a partial ionization, associate solutions, silicate slags). Particular attention was paid to constructing the models suitable for describing ordering and order/disorder transformations. 3. A thermodynamic assessment of numerous binary, ternary and multicomponent systems (including a usage of the results of first-principles calculations during the optimization). Compilation of the assessment results into specialized (*i.e.* material-oriented) databases. 4. A continuous improvement of program-user interfaces through making them more flexible and friendly.

My personal involvement in these "traditional CALPHAD activities" resulted in full or partial assessments of the Bi–Zn [3], Bi–Sn–Zn [4], Cu–Ni–Sn [5], Zn–Fe [6], Al–Fe–Mn–Si [7] and Zn–Fe–Al [8] systems as well as in proposing a method to establish confidence intervals of calculated phase boundaries [9]. This firsthand engagement with the thermodynamic optimization and the programs to carry it out (BINGSS/TERGSS [10] and PARROT) allowed me to identify a facet of the CALPHAD method that were not receiving a due attention during the three decades of developments and improvements – the assessment technique. A lack of attention resulted in numerous post-optimization phantoms, *i.e.* in calculated phase diagrams with aphysical artifacts such as high-temperature inverted miscibility gaps in the liquid phase [11,12]. It is worth realizing that the optimization module PARROT (which today is overwhelmingly used for assessments), BINGSS/TERGSS and other programs tailored for this purpose employ an unconstrained minimization. In the case of PARROT, it is known for sure that it "…makes use of the Powell method, available as subroutine VA05A in the Harwell Subroutine Library" [13]. According to the description of this subroutine [14], VA05 merely minimizes the sum of squares without the use of any partial derivatives, i.e. it is not apposite for a constrained optimization.

In the beginning of 2003, I was convinced that a constrained optimization had to become an integral part of the assessment technique. That belief was reflected in my previous application for a Discovery Grant. I am happy that my ideas about a topologically-constrained optimization have materialized in two published [15,16] and one accepted paper [17], as well as in two oral presentations at the 35th and 36th CALPHAD conferences.

The objectives of the research program

The main objective is to develop a new method of thermodynamic assessment, which takes into account a murky statistical nature of input data on the thermodynamic properties of phases, heterogeneous mixtures and conditions of phase equilibria. To reveal the essence of the problem to be tackled, let us realize that despite a thermodynamic flavour, from mathematical and computational angles the assessment of a particular system is a solution of either unconstrained or constrained non-linear least squares problem. In both cases, model parameters receive their statistically best numerical values accompanied by their covariance matrix. If this outcome is to be trusted and used, then it is required that input data possess certain statistical characteristics. In particular, it is necessary to ascribe a reasonable experimental error to each measurement, and to ensure that its connotation conforms to that adopted in mathematical statistics. Keeping these considerations in mind, let us recall the following.

1. In contrast to papers (especially in German journals) published in the XX century (especially in its first half), statistical characteristics of experimentally measured quantities are not paid due attention by a majority of authors today. A manifestation of this is that in contemporary publications tables of experimental data are more and more frequently replaced with figures visualizing these data. It is not a problem to scan and digitize these illustrations; the problem is that scanning and digitizing lead to quantities striped of their statistical meaning. 2. Important information crucial for pre-optimization analysis of experimental data is usually missing in modern publications. Instead, a painstaking analysis of experimental errors is often replaced with their estimations reflecting a personal opinion and experience of an expert proposing these estimations. This approach leads to the "experimental errors" without a statistical meaning. 3. Even if error bars are specified, it is seldom explained clearly and unambiguously what exactly is hidden behind them.

Subsequently, a rigorous analysis of the solution obtained is unattainable. Instead, numerous illustrations of proximity of calculated entities to their experimentally measured counterparts are given without attempting to quantify what propinquity is.

It is not intended to commence a hopeless struggle against a reproachable way in which experimental knowledge is disseminated today. It is contemplated to augment the CALPHAD method by a new assessment technique based on the ideas and algorithms of a non-statistical treatment of experimental information formulated by Leonid V. Kantorovich [18] and successfully employed for solving economic problems (Kantorovich won the Nobel Prize in economics in 1975) as well as for calculating enthalpies and entropies of sublimation and evaporation from tensimetric data [19-21]. Methodologically, this new technique will much better correspond to a fuzzy statistical nature of input experimental data than the habitual CALPHAD procedure. In plain English, I want to investigate whether the Kantorovich method can be employed for constructing the models of the Gibbs energies of phases from an array of experimental observations related to these phases.

The key idea is that every measurement can be characterized by its ultimate non-statistical error. By introducing the “interval calculus”, one ceases dealing with number of degrees of freedom, with the Student’s distribution, with the “three sigma” rule and so on. Instead, one creates a non-statistical corridor of errors associated with experimental data and demands that any function representing the observation must reside within it. Apparently, there is a region (not necessarily convex) within the space formed by models’ parameters any point of which satisfies the “within the corridor” condition. Since all points belonging to this region are equally acceptable, one can employ other criteria (*e.g.* an absence of inverted miscibility gaps or false inflection points on phase boundaries) for arriving at the most thermodynamically feasible solution.

The short-term objectives: 1. To use a meticulous analysis of various methods used in experimental thermodynamics and their specific features for finding consistent, unambiguous and explicit rules of assigning non-statistical ultimate errors to each basic experimental technique and its particular implementations. 2. To understand how a multidimensional region (not necessarily having a simple shape) of model parameters compatible with particular experiments can be constructed and to algorithmize this comprehension. To understand how the values of ultimate errors affect the size and shape of this region. 3. To understand how two or more regions are located if the experiment is related to conditions of phase equilibria or thermodynamic properties of multiphase mixtures when the properties of the phases are not independent. 4. To apply Kantorovich ideas to the thermodynamic assessment. To develop an algorithm of non-statistical optimization, implement it in a program and try it to assess previously optimized simple binary systems.

<u>The long-term objectives: </u>1. To compare the results of traditional and proposed types of optimization for several binary systems (it is very likely that the method proposed will first be tested on the Mg–Sb, Sn–Zr and As–Se systems) and a couple of ternary systems. To find similarities between these approaches. To identify irreconcilable differences and reveal methodological roots of these contradictions. 2. To intrigue the CALPHAD community with the new approach to the thermodynamic optimization (assuming that its usefulness and computational implementability would have been demonstrated) via publications and presentations. 3. To develop a rather general program for nonparametric optimization and disseminate it (for free, indeed!) To use feedback for refining the approach and writing concluding papers containing mathematical, numerical and thermodynamic aspects of the new method as well as examples of its applications in practice.

Methods and proposed approach

Let us consider a problem millions of scientists encountered millions of times. An experimental condition, x, is fixed, and a property, y, is measured. If this is done n times, one has a table x_i, y_i, Δy_i, $i = 1, \ldots, n$. If it is believed that the property depends on the condition, *i.e.* that $y = y\left(x, \vec{C}\right)$, then the table can be used for finding “optimal numerical values” of $\vec{C}$, which are named either “unknown coefficients” if $y\left(x, \vec{C}\right)$ is a convenient mathematical expression or “model parameters” if it is a physically feasible model. The “best values” of $\vec{C}$ are usually found by solving a corresponding least-squares problem $\sum_{i=1}^{n}\left(\left(y_i - y\left(x_i, \vec{C}\right)\right)\big/\Delta y_i\right)^2 \to \min\left(\vec{C}\right)$. Many important questions can be asked with respect to this procedure. Is the choice of model or formalism reasonable? Is the solution statistically acceptable? Is the solution physically sound? Now let us ask a question, which seems much simpler. What is Δy_i? Intuitively, it is clear that it reflects an accuracy of a measurement carried out for the experimental condition x_i. Quantitatively, this means that one has to acquire y_{ij}, $j = 1, \ldots, n_i$, *i.e.* results of repetitive measurements, from which Δy_i is calculated as $\sqrt{\sum_{j=1}^{n_i}\left(\bar{y}_i - y_{ij}\right)^2 \Big/ (n-1)}$.

Let us see how this works in practice. The weight of a penny, W, coined in 2007 was measured seven times using a Mettler PM1200 balance. In all seven instances, the weight was equal to 2.379 g, which resulted in useless $\Delta W = 0$. An inspection revealed a tag attached to the balance saying that $d = 0.001\,\mathrm{g}$. Although a precise meaning of d was not disclosed in the balance manual, one could assume that the weight could not be determined with the accuracy better than one milligram, *i.e.* that $\Delta W = 0.001\,\mathrm{g}$. One could also conclude that the Gaussian noise was less than $0.001\,\mathrm{g}$. Let us realize that ΔW associated with the value shown on the tag has a murky statistical nature. In other words, if the weight of the one-cent coin is given as $2.379 \pm 0.001\,\mathrm{g}$, $0.001\,\mathrm{g}$ should not be identified with the variance, it must be identified with a non-statistical confidence that an error of weight determination never exceeds one milligram. In fact, this means that the weight of the penny is fully characterized by the inequality $\underline{W} \le W \le \overline{W}$, where the lower limit $\underline{W} = 2.378\,\mathrm{g}$, and the upper limit $\overline{W} = 2.380\,\mathrm{g}$. The interval $\left[\underline{W}, \overline{W}\right]$ is not somewhat fixed once and forever. Imagine, for instance, that the Mettler balance is placed on a shaky table near an open window. Instead of calculating a new value of ΔW by employing repetitive measurements, one can widen the interval in such a manner that it would contain a correct value of weight for the new experimental environment (a wobbly table near an ajar window).

In the example considered, repetitive measurements were possible, but in thermodynamic investigations there are numerous situations when a measurement cannot be repeated. Imagine, for instance, that the enthalpy of mixing is being determined by drop calorimetry. It is very difficult to use exactly the same initial amount of the melt and to drop exactly the same amount of solid substance in two subsequent runs. If one starts not with unary melt and solid but with liquid and solid solutions, it is virtually impossible to have identical experimental conditions in two subsequent runs. One has a similar situation when DSC is employed for determining the enthalpy of a transformation. During a heating or cooling cycle, it is impossible in principle to measure a calorimetric signal (a difference in temperatures or a difference in powers) several times because an experimental condition is continuously changing. In both cases, however, a measured value of heat flux, $\dot{Q}_i$, can be characterized by a non-statistical interval $\left[\underline{\dot{Q}}_i, \overline{\dot{Q}}_i\right]$, which contains $\dot{Q}_i$ with the probability equal to unity. Although a width of the interval depends on quality of a device used for measurements as well as on skills and experience of an experimenter, there is no doubt in its existence.

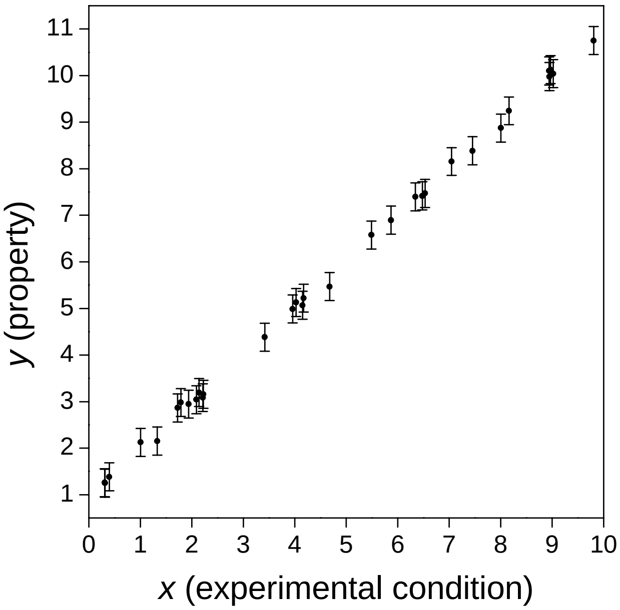

Figure 1

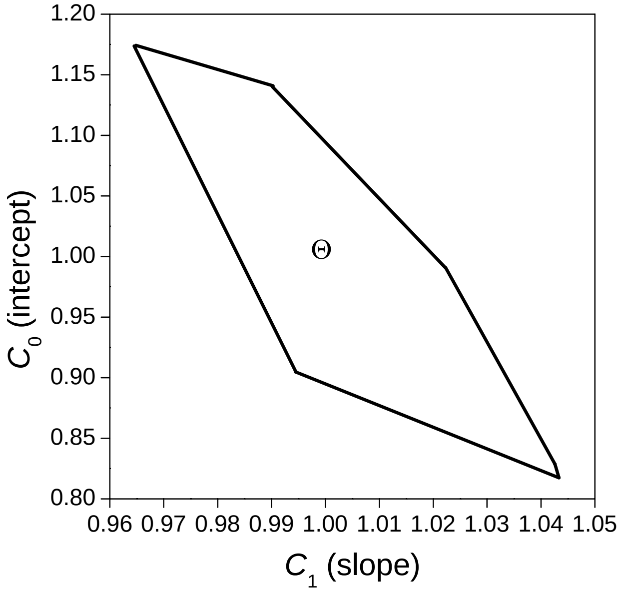

Figure 2

Let us see how the approach based on “interval calculus” can be used for extracting model parameters from experimental observations. Figure 1 shows results of an imaginary experiment. The error bars are not traditional ones! In our case, each error bar specifies the lower and upper limit associated with each measured property. By introducing the interval $\left[\underline{y}_i, \overline{y}_i\right]$ one guarantees that for the given experimental condition x_i, a measured value will always be confined within it. Data points strongly suggest that there is a linear dependency between y and x, which allows choosing $y(x,\vec{C}) = C_0 + C_1 x$. Since the error bars are non-statistical entities, a utilization of the least-squares method is unwarranted. Instead of seeking for the best statistical estimations of C_0 and C_1, let us find such a region of their values that any point belonging to it is in agreement with the experimental observations. This region Θ is shown in Figure 2. For an arbitrary point $\left(\tilde{C}_0, \tilde{C}_1\right)$ taken from this region, it is guaranteed that the straight line $\tilde{C}_0 + \tilde{C}_1 x$ will reside within the corridor formed by the error bars. In contrast to a classical treatment of the problem, the intercept and slope are characterized not by their best estimations and the covariance matrix, but by a peculiarly shaped region Θ. It is clearly seen from Figure 2 that the model’s parameters strongly correlate. It is worth noticing that Θ characterizes the correlation differently in comparison with a habitual covariation ellipse.

What if the same phenomenon was investigated not once, but two or more times? Figure 3 exemplifies a case of three independent examinations, A, B and C. It is worth accentuating that all error bars are the non-statistical intervals of “absolute trust” assigned to each measurement. At first glance, the results of all three experiments are in a good agreement, but Figure 4 reveals several important things. The results of experiments A and C are in an agreement, because Θ_A and Θ_C overlap. The fact that $\Theta_\mathrm{B} \cap \Theta_\mathrm{A} = \varnothing$ and $\Theta_\mathrm{B} \cap \Theta_\mathrm{C} = \varnothing$ does not necessarily mean that the results of the experiment B are erroneous; it means that the outcomes of all three experiments are to be critically analyzed again and that extra experimental efforts may be needed.

The considerations and examples presented here were deliberately made very simple to clarify some important distinctive features of the research methodology. A much deeper mathematical analysis that can be found in [18,22] will form an actual methodological basis for this research.

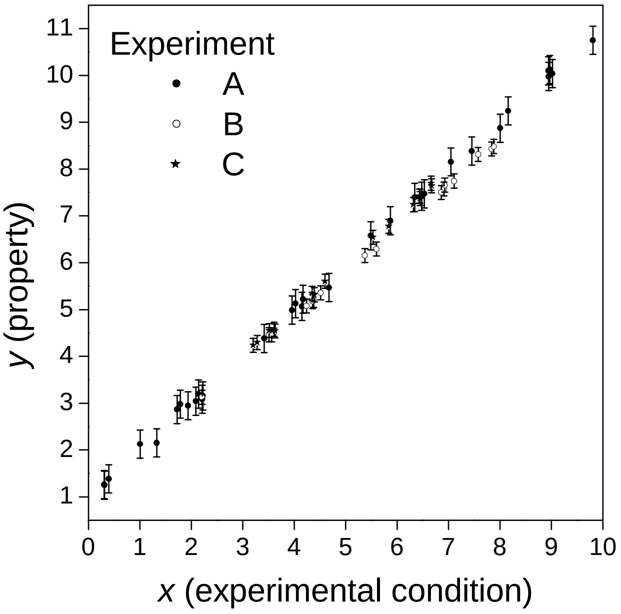

Figure 3

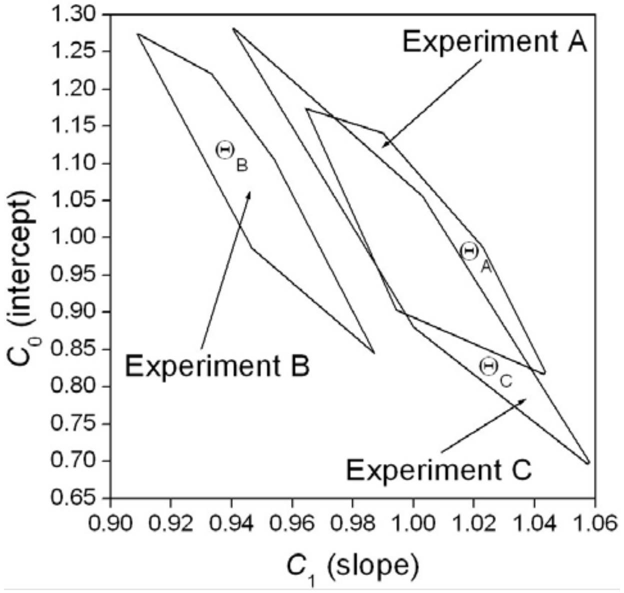

Figure 4

Anticipated significance of the work

Hundreds if not thousands of researchers all over the world use the CALPHAD method for constructing the Gibbs energies of phases via a simultaneous treatment of heterogeneous experimental data. These models are then employed for solving a wide variety of scientifically interesting and industrially important problems [23]. Whether the solutions of these problems are trustworthy or not depend (among many other things, indeed) on an internal self-consistency of a procedure chosen for data treatment. It will be advantageous for the CALPHAD community if in addition to the traditional assessment technique it has a new method whose validity is not jeopardized by a statistical murkiness of input data and which, in principle, allows a multicriterial optimization, during which inverted miscibility gaps, false inflection points and other unwanted artifacts are being suppressed in the course of assessment.

Training to take place through the proposal

It is clear from the Budget that NSERC funding will be used to support only one Doctoral student, which is not ambitious, indeed. The reason is that the project is tough and risky. Let us realize that it is intended to apply a non-traditional and mathematically sophisticated method to thermodynamic optimization, i.e. to use it in the area it has never been employed before. Moreover, nobody has ever thought about whether the method is applicable in principle in this area or not. Subsequently, there is no previous knowledge we can lean against. By saying this I do not mean that the project is hopeless, I do not look for excuses of a possible failure in advance. What I am saying is that this endeavor will require dedication and a great deal of attention, that shall I have to figure out many things by myself, that it will be painful to find a right direction and to make the first steps. Under these circumstances, I cannot guarantee that I shall be an efficient supervisor for two or more students. I can be a mentor for only one hand-picked individual. If it is realized that the project leads to nowhere, then the direction of student's research will be re-oriented for ensuring that she/he will defend her/his thesis in time.

References

1. N. Saunders, A.P. Miodownik "CALPHAD, calculation of phase diagrams: a comprehensive guide" Pergamon, 1998
2. H.L. Lukas, S.G. Fries, B. Sundman "Computational thermodynamics: the CALPAD method" Cambridge University Press, 2007
3. D.V. Malakhov "Thermodynamic assessment of the Bi–Zn system" *CALPHAD*, 2000, Vol.24, No.1, P.1-14
4. D.V. Malakhov, X. J. Liu, I. Ohnuma, K. Ishida "Thermodynamic calculation of phase equilibria of the Bi–Sn–Zn System" *Journal of Phase Equilibria*, 2000, Vol.21, No.6, P.514-520
5. G.R. Purdy, D.V. Malakhov, A. Guha "Homogenization of multicomponent alloys via partial melting" *Journal of Phase Equilibria*, 2001, Vol.22, No.4, P.439-450
6. J. Nakano, D.V. Malakhov, G.R. Purdy "A crystallographically consistent optimization of the Zn–Fe system" *CALPHAD*, 2005, Vol.29, No.4, P.276-288
7. M.K. Abou Khatwa, D.V. Malakhov "On the thermodynamic stability of intermetallic phases in the AA6111 aluminum alloy" *CALPHAD*, 2006, Vol.30, No.2, P.159-170
8. J. Nakano, D.V. Malakhov, S. Yamaguchi, G.R. Purdy "A full thermodynamic optimization of the Zn–Fe–Al system within the 420–500ºC temperature range" *CALPHAD*, 2007, Vol.31, P.125-140\
9. D.V. Malakhov "Confidence intervals of calculated phase boundaries" *CALPHAD*, 1997, Vol.21, No.3, P.391-400

10. H.L. Lukas, S.G. Fries “Demonstration of the use of BINGSS with the Mg–Zn system as example” *Journal of Phase Equilibria*, 1992, Vol.13, No.5, P.532-541
11. Y.A. Chang et al “Phase diagram calculation: past, present and future” *Progress in Materials Science*, 2004, Vol.49, P.313-345
12. R. Schmid-Fetzer et al. “Assessment techniques, database design and software facilities for thermodynamics and diffusion” *CALPHAD*, 2007, Vol.31, No.1, P.38-52.
13. B. Jansson “Computer operated method for equilibrium calculations and evaluation of thermochemical model parameters” Ph.D. thesis, Royal Institute of Technology, Stockholm, 1984
14. http://hsl.rl.ac.uk/archive/hslarchive.html
15. D.V. Malakhov, T. Balakumar “Re-optimization of the Mg–Sb system under topological constraints” *International Journal of Materials Research*, 2006, Vol.97, No.5, P.517-525
16. D.V. Malakhov, T. Balakumar “Post-optimization elimination of inverted miscibility gaps” *International Journal of Materials Research*, 2007, Vol.98, No.9, P.786-796
17. Dmitri V. Malakhov, Thevika Balakumar “Undulate phase boundaries on binary T-x diagrams” (accepted for publication in *CALPHAD*)
18. L.V. Kantorovih “On new approaches to computational methods of data treatment” Siberian Mathematical Journal, 1962, Vol.3, No.5, P.701-709
19. G.A. Kokovin, V.A. Titov, A.A. Titov, S.I. Spivak “Some methodological aspects of mathematical treatment of experimental data on equilibrium” In: *Mathematics in chemical thermodynamics*, Nauka, Novosibirsk, 1980, P.50-58
20. E.F. Titova, G.A. Kokovin “On an applicability of the non-statistical approach to dual non-linear problems” In: *Mathematical methods in chemical thermodynamics*, Nauka, Novosibirsk, 1982, P.170-177
21. E.F. Titova, G.A. Kokovin, T.P. Chusova, V.A. Titov “On a simultaneous treatment of static and dynamic tensimetric data” In: *Informal mathematical models in chemical thermodynamics*, Nauka, Novosibirsk, 1991, P.81-92
22. V. Kreinovich, A.V. Lakeyev, J. Rohn, P.T. Kahl “Computational complexity and feasibility of data processing and interval computations” Springer, 1997
23. K. Hack (editor) “The SGTE casebook: thermodynamics at work” London: Institute of Materials, 1996