

This article was downloaded by:

On: 14 January 2011

Access details: *Access Details: Free Access*

Publisher *Taylor & Francis*

Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



## **Molecular Simulation**

Publication details, including instructions for authors and subscription information:

<http://www.informaworld.com/smpp/title~content=t713644482>

## **Inverse Problem in Materials Science**

E. B. Rudnyi<sup>a</sup>; V. V. Kuzmenko<sup>a</sup>; G. F. Voronin<sup>a</sup>

<sup>a</sup> Chemistry Department, Moscow State University, Moscow, Russia

**To cite this Article** Rudnyi, E. B. , Kuzmenko, V. V. and Voronin, G. F.(2000) 'Inverse Problem in Materials Science', *Molecular Simulation*, 24: 1, 191 — 196

**To link to this Article:** DOI: 10.1080/08927020008024195

**URL:** <http://dx.doi.org/10.1080/08927020008024195>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

## INVERSE PROBLEM IN MATERIALS SCIENCE\*

E. B. RUDNYI<sup>†</sup>, V. V. KUZMENKO and G. F. VORONIN

*Chemistry Department, Moscow State University, 199899 Moscow, Russia*

*(Received April 1999; accepted May 1999)*

The importance of the inverse problem in materials science, that is, the determination of unknown parameters in the physico-chemical models from the experimental values is emphasized. The main attention is given to the problem of how to take into account experimental systematic errors during the inverse problem.

**Keywords:** Inverse problem; thermodynamic assessment; systematic errors; linear error model

### INTRODUCTION

The advances of material science depend heavily on reliable thermodynamic databases which allow us to optimize the synthesis conditions for the given material and to predict its behavior in different mediums.

Thermodynamic properties required, in principle, can be obtained during two steps from first principles. *Ab initio* calculations give us the potential surface for the system in question, and then statistical thermodynamics or molecular dynamics brings forward the thermodynamic functions. In practice, these two steps are separated from each other, and quite often for the second step the interaction potential is taken not from *ab initio* calculations but rather from semi-empirical approach (for example, molecular mechanics).

Both tasks are usually referred to as direct problems. That is, there is some mathematical problem to solve (Schrödinger equation, configuration

---

\* Presented at Molecular Simulation '99 Electronic Conference.

<sup>†</sup>Corresponding author. e-mail: rudnyi@comp.chem.msu.su

integral or motion equation), and all the parameters within it are assumed to be known.

However if we speak about quantitative results, the precision of non-empirical calculations still remains quite unsatisfactory. When the solution of the real practical problem in material science is required, no one would bet only on the results obtained by the approaches described above, because the precision of the experimental measurements available is by orders of magnitude higher than those from non-empirical calculations. On the other side, the experimental values by themselves do not give us the full answer either. What is necessary here is multi-dimensional interpolation or even extrapolation, and in order to be reliable this should be based on some physical model.

Thus, the only reasonable way to proceed is to switch to an inverse problem. Here the results obtained by computational chemistry are treated as qualitative. They give us a clue what the functional form may look like, but the parameters within it are assumed to be unknown and they should be determined from the experimental results available. Along this way, we can achieve the description of Nature within the experimental precision which currently is rather high and at the same time to predict the behavior of the system at the conditions where the experiments can not be carried out.

In modern material science the CALPHAD group (see papers in the journal under the same name) presents rather successful example of employing the inverse problem for numerous tasks in metallurgy, ceramics, geology, semiconductors, high-temperature superconductors and other areas. There are several integrated commercial packages available as well as public domain and shareware software (for example, see lists [1,2]) based on the rather reliable databases of thermodynamic properties obtained during the inverse problem – a simultaneous assessment of the Gibbs energy of phases from heterogeneous experimental values (thermodynamic properties and phase diagrams).

The main stress in the current work is given to one general problem arising in the inverse problem: taking into account systematic errors which are inevitable in real experimental measurements. The practical examples of the successful use of the approach suggested are given.

## FORMAL TASK

The inverse problem starts with a set of experimental values obtained by different authors and by different experimental methods. For most cases,

they can be expressed as follows

$$\{y_{ij}, x_{ij}, z_i\} \quad (1)$$

- $i$  enumerates heterogeneous experiments,  $i = 1, \dots, N_i$ ,
- $j$  enumerates the experimental points in the  $i$ -th experiment,
- $y_{ij}$  is output,
- $x_{ij}$  is input,
- $z_i$  is the vector of controlled values that were constant in the  $i$ -th experiment.

Note that even when an experiment is multidimensional, the experimenter usually changes a single variable during a single run.

It is assumed that there is some model that can describe all the experimental results

$$y_{ij} = f_i(x_{ij}, z_i; \Theta) + \varepsilon_{ij} \quad (2)$$

$f_i$  is a relationship between input and output in the given  $i$ -th experiment,  $\Theta$  is a vector composed from the parameters to be determined, and  $\varepsilon_{ij}$  is the measurement error. Note that  $f_i$  may be different but all of them depend on the same unknown parameters.  $f_i$  may be given in the closed form or as some numerical task.

Formally speaking, the task of the simultaneous assessment is to obtain a set of unknown parameters in Eq. (2) that gives the best description of the original experimental values. The main problem lies in question what should be considered as the best description of the experimental points. Actually the number of unknowns in the system (2) is always greater than the number of equations because experimental errors  $\varepsilon_{ij}$  are also unknown. Therefore, there is an infinite number of solutions and which one should be taken as the best strongly depends on our considerations of errors. The latter will be referred to as the error model and should not be confused with the physico-chemical model taken by itself.

The conventional approach is to employ the weighted least squares method, *i.e.*, to find such a solution that brings the sum of squares of the errors

$$SS = \sum_{ij} \varepsilon_{ij}^2 W_{ij} = \varepsilon' W \varepsilon \quad (3)$$

to the minimum. In matrix notation  $\varepsilon$  is the vector that comprises all the errors  $\varepsilon_{ij}$  from all the experiments (the number of elements  $\sum_i N_i$ ) and  $W$  is

the weight matrix that contains weights for each experimental point on its diagonal,  $\mathbf{W} = \text{diag}\{W_{ij}\}$ .

The problem that has got no clear answer in the weighted least squares is how to assess the weights. The final solution certainly depends on the values of the accepted weights and a different set of weights would lead us to the different solution. Consequently, in the weighted least squares method the task of the simultaneous assessment becomes mainly of that of the weight assignment. Mathematical statistics gives a guideline such that in order to obtain the reliable solution the weight matrix should be equal to inverse of the dispersion matrix of the error vector up to the factor

$$\mathbf{W} = k\mathbf{D}(\boldsymbol{\varepsilon})^{-1} \quad (4)$$

We can proceed from Eq. (4) to the weight least squares if all the errors  $\varepsilon_{ij}$  are postulated to be non-correlated (only in this case the dispersion matrix takes the diagonal form) and ratios between variances for all the errors are known *a priori*. Unfortunately, both statements are too restrictive for real-life applications. The variance ratio for experimental points is not known and there are clear evidences that at least some errors are correlated between each other because of the systematic errors. Then let us start with Eq. (4) and develop a more general approach than the weight least squares.

If variances are not known the maximum likelihood method allows us to determine both unknown parameters (vector  $\boldsymbol{\varepsilon}$ ) and variances simultaneously by means of maximizing the likelihood function. This permits us to drop the requirement for variance ratios to be known. Provided all the errors are described by the multinormal distribution the maximum of the likelihood function coincides with the maximum of

$$L = -\ln\{\det[\mathbf{D}(\boldsymbol{\varepsilon})]\} - \boldsymbol{\varepsilon}'\mathbf{D}(\boldsymbol{\varepsilon})^{-1}\boldsymbol{\varepsilon} \quad (5)$$

Note that the weighted least squares method is a special case of maximizing Eq. (5) when the variance components are known up to a constant, that is the maximum of (5) matches the minimum of (3) provided there are no other unknowns inside the dispersion matrix but the general factor (Eq. (4)).

The linear error model

$$\varepsilon_{ij} = \varepsilon_{r,ij} + \varepsilon_{a,i} + \varepsilon_{b,i}(x_{ij} - x_i) \quad (6)$$

- $\varepsilon_{ij}$  is the total error,
- $\varepsilon_{r,ij}$  is the reproducibility error (experimental noise),
- $\varepsilon_{a,i}$  is the shift systematic error (shift laboratory factor),
- $\varepsilon_{b,i}$  is the tilt systematic error (tilt laboratory factor),
- $x_i$  is a mean of  $x_{ij}$  in the  $i$ -th experiment.

has been recently introduced [3]. It is assumed that the total experimental error  $\varepsilon_{ij}$  consists not only from the reproducibility error  $\varepsilon_{r,ij}$  but also from two systematic errors  $\varepsilon_{a,i}$  and  $\varepsilon_{b,i}$ . Both systematic errors are constant within the  $i$ -th experiment but they are assumed to change randomly among different experiments. The former systematic error accounts for the shift systematic error and latter for the tilt laboratory factor (tilt systematic error). Note that the linear error is a special case of so-called mixed models (see [4]).

The practical reason for introducing new two terms in Eq. (6) is that the results of the distinct experiments usually differ more between each other than the reproducibility error in a single experiment. Formally speaking, there is a statistically significant difference between distinct experiments, *i.e.*, the ratio of the corresponding sum of squares is more than Fisher's criterion allows. Then the systematic errors introduced above permit us to treat this situation by means of formal statistical procedures.

Equation (6) was designed for one-dimensional tasks. Fortunately, it can be applied for most practical problems because, as was mentioned above, under the conventional approach only one variable is changed within a single run. Thus all the experiments can be usually treated as pseudo-one-dimensional. As experimental material science switches to multidimensional experimental design (modern analytical chemistry appears to be doing so), the linear error model may need to be modified.

The linear error model brings forth the dispersion matrix of experimental errors,  $\mathbf{D}(\varepsilon)$  in the block-diagonal form (see details in [3]). Each block corresponds to the single experiment and its elements are functions of three variance components

$$D(\varepsilon_{r,ij}) = \sigma_{r,i}^2, \quad D(\varepsilon_{a,i}) = \sigma_{a,i}^2, \quad D(\varepsilon_{b,i}) = \sigma_{b,i}^2. \quad (7)$$

The considerations above allow us to set up a task as follows. For the given experimental points (1), it is necessary to determine vector  $\Theta$  with unknown parameters in the physico-chemical model and unknown variance components contained in the dispersion matrix simultaneously. The maximum likelihood method provides a framework to achieve this goal and also the

criterion for the best solution for the system (2). The algorithm for maximizing Eq. (5) under the linear error model is described in Ref. [3].

### Online Examples (<http://www.chem.msu.su/~rudnyi/>)

- Combined Processing of Experimental Series with Systematic Errors – Non-Linear Physico-Chemical Model with Linear Error Model (see also [3]).
- Simultaneous Assessment of Thermodynamic Properties of Calcium Aluminates (see also [5]).
- Thermodynamic Assessment of the Ba-Cu-Y System (see also [6]).
- Simultaneous assessment of  $\text{YBa}_2\text{Cu}_3\text{O}_{6+z}$  phase thermodynamics (see also [7]).

### References

- [1] Gordon, T. M., “Software and Data for Thermodynamics and Phase Equilibrium Calculations in Geology”, <http://ichor.geo.ucalgary.ca/~tmg/Research/thermolinks.html>.
- [2] Bale, C. W., “Web Sites in Inorganic Chemical Thermodynamics”, <http://www.crct.polymtl.ca/fact/websites.html>.
- [3] Rudnyi, E. B. (1996). “Statistical model of systematic errors: linear error model”, *Chemometrics and Intelligent Laboratory Systems*, **34**(1), 41–54.
- [4] Rao, C. R. and Kleffe, J. (1988). Estimation of variance components and applications, North-Holland, Amsterdam, *North-Holland Series in Statistics and Probability*, **3**, 370.
- [5] Kuzmenko, V. V., Uspenskaya, I. A. and Rudnyi, E. B. (1997). “Simultaneous assessment of thermodynamic functions of calcium aluminates”, *Bulletin des Societes Chimiques Belges*, **106**(5), 235–243.
- [6] Rudnyi, E. B. (1997). “Statistical model of systematic errors: An assessment of the Ba-Cu and Cu-Y phase diagram”, *Chemometrics and Intelligent Laboratory Systems*, **36**, 213–227.
- [7] Rudnyi, E. B., Kuzmenko, V. V. and Voronin, G. V. (1988). “Simultaneous assessment of the  $\text{YBa}_2\text{Cu}_3\text{O}_{6+z}$  thermodynamics under the linear error model”, *J. Phys. Chem. Ref. Data*, **27**(5), 855–888.