Assessment module in OC

Bo Sundman January 7, 2016

VERY PELIMINARY DRAFT

Contents

1 Background

Thermodynamic software is useless without validated thermodynamical databases. The work to develop databases of high quality is tedious and slow. It is also lmited by the thermodynamic models implemented in the software. Thus it is important that the software can take advanateg of new experimental and theoretical data for the development of new models. The problem with proprietary software for thermodynamic calculations is that the investment in good databases makes it less interesting to develop new models becaus that may render existing databases less important. The development of databases and new modelling software must go hand in hand. This is one of the main resons to develop a free thermodynamic software like Open Calphad.

The OC software is available at [http://www.opencalphad.org] or as the opencalphad repository at [http://github.com].

A good introduction to thermodynamic models and the assessment procedure can be found in the book by Lukas et al[07Luk].

1.1 Modeling crystalline phases

For crystalline phase the Compound Energy Formalism (CEF)[01Hil] is almost universally accepted within the Calphad community. This can handle long range ordering (LRO) and multiple sublattices with different kinds of species. Thus interstitial solutions, defects models, carbides, oxides with ionic constiutuents etc can be described accurately with CEF based models. CEF cannot describe explicitly short range ordering (SRO) but with few exceptions the contribution to the Gibbs energy due to SRO can be modelled with CEF[07Luk]

The Gibbs energy can be partitioned for phases with long range ordering (sublattices) in an ordered and a substitutional part. The implementation of this for phases with order/disorder transformations, like $A1/L1_2/L1_0$ has been revised and simplified.

In commercial software ordering in FCC and BCC phases have for a long time been limited to a tw-sublattice model even if a 4 sublattice model can better describe the different types of ordering. One reason has been that the implementation of the 4 sublattice model has been very slow and industrial users find thus cumbersome. But with better implementations and using parallelization of the software such problems can be solved at least partially.

Intermetallic phases like σ, μ and others have also been simplified in existing databases using fewer sublattices that exists crystallographically. Such problems can also be solved by a new and better thermodynamic software.

1.2 Modeling liquids

For the liquid phase there is no really good model available. In many case a regular solution model is sufficent but when there is SRO in the liquid phase there are many different models available but none of them is really good. See the discussion in [07Luk].

The modified quasichemical model, strongly argued by the FactSage group, is unphysical in the sense that it assumes 2 bonds per atom in the liquid in order to avoid negative configurational entropy. This also means that positive interactions between the constituents of the liquid does nor result in a miscibility gap, such miscibility gaps must be modelled with additional interaction parameters. The model may describe experimental data accurately but it lacks a sound scientific basis.

The associated model assumes covalent bondig between the atom forming molecules (called associates) in the liquid and this leads to some problems modelling the properties close to the pure elements when the associates when the liquid is more or less regular are not really important. For multicomponent systems it also adds the problem to include interactions between the associates and additional constituents. For ionic systems the assumtion of covalent bonding forming associated is not in accordance with the reality which must include some the carge transfer.

The partially ionic 2 sublattice liquid model (I2SL) assumes the SRO is mainly due to charge transfer and that the liquid can be described as LRO where cations and anions do not mix. To handle systems without anions (metallic liquids) it assumes there are charged vacancies in a "valency band" and for liquids with mainly anions these are assumed to be exist also as neutrals. For binary systems the I2SL model is more or less equivalent to the associated model (using a single associate) and for multicomponent systems it requires fewer interaction parameters than the associated model. The main problem with the I2SL model is its strong tendency to predict reciprical miscibility gaps which can be difficult to control.

There is a proposal for new liquid model, called the "corrected quasichemical model[09Hil]" which can describe SRO using a quasichemical model with a realistic number of bonds between the constituents of the liquid and which correctly predics miscibility gaps when there is a prositive interactions without any extra regular solution parameters. However no commercial software has implemented this model because if is is successful it would require reassessments of several large databases.

This means there is a need for a free scientific software to push the development of thermodynamic models in order to improve the databases.

2 Equilibrium calculations

The equilibrium calculations in OC follows the algorithm proposed by Hillert [81Hil] and explained in more detail by Sundman et al[15Sun2] and in the documentation of the HMS package. It allows a very flexible set of conditions on amounts or fractions of components,

chemical potentials or activites, enthalpies, fix phases etc.

2.1   Parallelization of equilibrium calculation

It will be possible to use parallelization during assessments in OC and this will grately speed up the assessment procdeure. Each experiment represent an equilibrium situation with one or more experimental information. With OC can calculate several equilibria with different sets of conditions in parallel using the OpenMP library.

2.2   Calculation of partial derivatives

During equilibrium calculations OC calculates all second derivatives with respect to T, P and all constituents of all phases. This improves stability and makes it possible to calculate for example the stability function as part of an equilibrium calculation.

It also makes it possible to implement The "dot derivative" method to calculate derivatives of state variables has been implemented. This allows calculation of properties like the heat capacity without resorting to numerical derivation. It makes use of the analytical first and second derivatives of the Gibbs energy for T, P and all constituent fractions implemented in the model package. The implementation is not complete but derivatives of several state variables with respect to T are available.

2.3   Calculation and plotting of diagrams

During an assessment one has to calculate many different kinds of diagrams because the experimental data are just a few points whereas the models describe the whole composition and temperature range. Many things can happen in the regions where there is little or no data and these must be taken into accound. For example an intermetallic phase stable only in a small temperature and composituon range may appear as stable outside this range. Even if there is no experimental data in this range it is most probably an error in the modelling that it appears and the person respnsible for the assessment must supress this by modifying the model parameters for this intermetallic phase. This can be done by adding fictitious experimental data prescribing that the driving force for the phase should be negative where the phase should not be stable.

One type of diagram is the property diagram either varying only one variable for example the T or a composition. This is generated by the STEP procedure in OC. There is also a "step separate" option for Gibbs energy curves and similar things when each phase is calculated separately along the axis.

The MAP procedure for phase diagrams calculates lines where the set of stable phases changes for different values of the axis variables. At present only two axis are allowed.

3 The beginning of an assessment procedure

The beginnings of an assessment procedure has been implemented is a straightforward least square fit of experimental and theoretical data to the same data calculated by the model by varying model parameters. The experimental data can be entered as equilibria and in fact very little had to be done in the OC software for this purpose. Oc make use of a free least square subroutine called LMDIF developed at Argonne National Lab 1980 and part of the MINPACK software.

3.1 The function to be minimized

The assessment prcedure tries to modify the model parameters selected by the assessor in order to minimize the function F:

$$F(q_i) \;\; = \;\; \sum_j (\frac{x_j^{\mathrm{exp}} - x_j^{\mathrm{calc}}(q_i)}{\sigma_j} w_j)^2 \tag{1}$$

where q_i represents the model parameters selected by the assessor to be varied, x_j^{exp} is the measured experimental propery j and $x_j^{\mathrm{calc}}(q_i)$ is the same property j calculated from the model using the parameters q_i. σ_j is the experimental uncertainties for the experimental property j and w_j is a weight assigned by the assessor to the experiment. It is possible also to have inequality experimental data if only a limiting value is known.

The LMDIF subroutine developed by Argonne [80Arg] has not really been tested in a full scale assessment but in some small tests it seems to have good stability and works well.

3.2 Subroutines in OC

In OC a single subroutine was added to the HMS pacakge which is called by LMDIF in order to calculate the terms in eq. 1 for each new set of model parameters set by the LMDIF package.

```
  subroutine assessment_calfun(nexp,nvcoeff,f,x)
! nexp is number of experiments, nvcoeff number of coefficients
! f is the differences between experiments and value calculated by model
! returned by this subroutine
! x are the current model parameter values set by LMDIF
    implicit none
    integer nexp,nvcoeff
    double precision F(*),X(*)
```

Some additional subroutines were also added in the GTP pacakage to allow this least square routine to change model parameters and to list results and handle experimental data. A number of new commands had to be added for the user to set up and control the assessment process, these are described in section 4.2.

What is severely missing are additional software to provide statistical analysis of the results like estimating uncertainties using the assessment results for extrapolations.

3.3 Assessment of other properties that depend on T, P and phase composition

OC has a flexible way to handle properties like mobilities, elastic constants etc that may depend on the phase, T, P and the phase constitution. Some properties are predefined but a skilled programmer can easily add a specific property and a model to use it in a calculation. The values of such properties can be obtained interactivly or by application software in the same way as thermodynamic state variables. They can also be assessed with the same procdeure as thermodynmaic model parameters.

4 Performing an assessments in OC

An assessment would roughly be performed in the following steps:

1. On a macro file file you have the commands to:

   - ENTER elements, species, phases, models, known model parameters etc. from a macro file.
   - ENTER OPT to enter the coefficients to be optimized, A00 to A99.
   - ENTER PARAMETERS with coefficients to be optimized using the symbols A00 to A99.

2. On the same or on another macro file you have the experimental data.

   - The experimental data is entered as part of an equilibrium using ENTER EQUILIBRIUM or ENTER MANY_EQUILIBRIA commands.
   - In addition to the necessary conditions to calculate the equilibrium you use the ENTER EXPERIMENT command to specify the experimental property and value.
   - With the command SET RANGE_EXP_EQUIL you specify the range of equilibria with experimental data.

3. The rest of the assessment is done interactivly and the set of commands are used as needed. These must normally be reiterated many many times using different weights of the equilibria and optimizing variables.

- You SET VARIABLE_COEFF to specify a coefficient to be optimized.
- You OPTIMIZE to make a least square fit using all experimental data except thos with weight zero.
- You LIST OPTIMIZATION to list the current result.
- You may SET WEIGHT on the different equilibria. Weight zero means the experiment will be ignored.
- You may AMEND ALL_OPTIM_COEFF to rescale or recover parameters.
- You may SELECT EQUILIBRIUM to select an equilibrium and use many other commands to change conditions, phase status etc for this equilibrium.

4. When you finished you can use SAVE TDB or LIST DATA TDB to create a TDB file with the results (it will need editing). At present there is no way to save your results inside the OC software. It is recommended to use a LOG file to document the way you obtained your results.

4.1 Assessment examples

There are two short maro examples how to set up an assessment and start the optimization. OC has not yet been used for a full assessment, you have the chance to be the first to do that and publish the result!

4.2 Special commands for an assessment

The special commands for performing an assessment is described a little more in detail here. Many other commands in OC can also be used as explained above.

ENTER OPTIMIZE_COEFF enters up to 100 coefficients to be optimized. They are called A00 to Aij (default A99). This must be done before any assessment. The symbols Aij can be used when entering TP functions or parameters.

ENTER EXPERIMENT the measured property of a state variabl e or symbol, its value and uncertainity is added to the current equilibrium with this command.

ENTER EQUILIBRIUM "name" experimental data are always related to an equilibrium. This command enters the necessary data structure to set to conditions for an experiment and to add experimental data. All conditions must be set explicitly for each equilibriuum.

ENTER MANY_EQUILIB to enter tables of experimental data. This command allows the user to enter a table with conditions for many equilibria with similar conditions. It is intended for assessments but can be used also to calculate various properties for a range of conditions not suitable for a STEP calculation. The user must give a general "head" of the table specifying phase status and all conditions including some that are different in the different equilibria. For those familiar with Thermo-Calc this command replaces tables in a POP file in assessments but is more flexible.

- This command will first ask Table head::
  and the user must specificy conditions and other things for the table. The user can refer to values (or text) in the table by giving an "@" character followed by a digit 1 to 9. The digit represent the column with the value. In the example entering a table row below the user input is **in bold** and explanatons in *italics*.

  By default all phases are suspended for the equilibria so the user must specify the phases that should be entered, dormant or fixed.
  Table head:: **entered 0 *** *all phases should be entered*
  Table head:: **fix 0 liquid** *liquid should be fix with 0 moles*
  Table head:: **fix 1 @2** *the phase in column 2 should be fix with 1 moles*

- The conditions can be given using the @ character to indicate vaules that are given in the specified column in table to follow.
  Table head:: **condition t=@1 p=1e5 n=1 w(cr)=@3 w(mo)=@4**

- Optional calculations of entered symbols
  Table head: **calculate cp**

- Optinonal listing of state variables
  Table head: **list HM tc(bcc)**

- Optional experimental data
  Table head: **experiment x(liquid,cr)=@5:.01, x(bcc,cr)=@6:.02**

- Optional comment
  Table head: **comment experimental data from Kubaschwki 1955**
  such comments will be listed with the equilibrium.

- Comment lines starting with "@$" can be inserted anywhere but will not be saved.

- The table head is finished by an empty line or "table_start"

- For the rows in the table the user must first provide a unique name for each equilibrium (that is counted as column 0 (zero)) and values for all columns referred to in the table head.
  Table row:**EQ1 1573 BCC 0.3 0.05 0.27 0.42**
  Table row:**EQ2 1623 BCC 0.35 0.07 0.28 0.51**

- The table is finished by
  Table row:**table_end**

AMEND ALL_OPTIM_COEFF to rescale or recover their values.

LIST OPTIMIZATION lists results of the optimization in various ways. Only the SHORT option implemented.

LIST DATA TDB lists all current parameters and assessed model parameters on a TDB file. Some editing are needed.

OPTIMIZE calls the optimizer to obtain a least square fit of the values calculated from the model parameters to the experimental data.

SET RANGE_EXP_EQUIL "first" "last" must be given to specify the equilibria with experimental data. They must have been entered sequentially. An AMEND RANGE will be added.

SET OPTIMIZING_COND to set some parameters for the optimizer.

SET VARIABLE_COEFF "index" "start value" set a start vaule of a coefficient ij to be optimized.

SET SCALED_COEFF "index" "start value" additionally asks for scaling factor, a min and max value for a coefficient. Not yet implemented.

SET FIXED_COEFF "index "value" to set an optimizing coefficent ij to a fixed value (usually its current value)

SET WEIGHT "value" "range" assigns a value to the weight (default unity) to equilibria with names fitting the "range". If the weight is set to zero the experiments from these equilibria will be ignored.

SELECT EQUILIBRIUM "name/number" to select an equilibrium to perform some changes.

There is no special "edit" module as in Thermo-Calc for modifying the equilibria with experiments, this can be done using the **select_equilibrium** command. The first equilibrium (DEFAULT) cannot be used for experimental data. Using **step** and **map** will also create new equilibria for node points, they may interfere with the experimental equilibria, this has to be tested. I have to implement some nice way to get rid of redundant equilibria, the **delete equilibria** is a bit dangerous. There is no way yet to save results like on a PAR file in Thermo-Calc.

References

[80Arg] Argonne National Lab (1980) the LMDIF package.

[81Hil] M Hillert, Physica, **103B** (1981) 31

[01Hil]   M Hillert, J of Alloys and Comp **320** (2001) 161

[07Luk]   H L Lukas, S G Fries and B Sundman, *Computational Thermodynamics*, Cambridge Univ Press (2007)

[09Hil]   M Hillert, M Selleby and B Sundman, An attempt to correct the quasichemical model, Acta Materialia, **57** (2009) 5237-5244

[http://www.opencalphad.org] Official Open Calphad web site

[http://github.com] opencalphad repository at http://github.com.

[15Sun1] B Sundman, U Kattner, M Palumbo and S G Fries, OpenCalphad - a free thermodynamic software, Integrating Materials and Manufacturing Innovation, **4**:1 (2015), open access

[15Sun2] B Sundman, X-G Lu, H Ohtani, Comp Mat. Sci **101**(2015) 127