Datafile construction for THERMOCALC 3.1 back

The information here relates to datafile construction for

mode 1: calculating phase diagram information for systems involving solid solutions

mode 2: calculating average PT estimates, and

mode 3: calculating phase diagrams for systems not involving any solid solutions.

Mode 0, which allows end-member information to be extracted from the pdata files and THERMOCALC itself, can use any (valid) mode 1–3 datafile.

The datafiles for mode 1 are by far the most complex, as in this mode THERMOCALC calculates not only the T of an equilibrium at given P, for example, but it also calculates the *compositions* of the phases involved in the equilibrium. This means that you have to communicate to THERMOCALC—via your datafile that the software reads—the activity-composition (a-x) relationships for the phases[1].

The same code is used in THERMOCALC to read all datafiles. However, the meaning of what is read may be different depending on the mode specified, the difference being obvious in the output. The mode is specified either as one of the first questions that THERMOCALC asks you when you run it, or it can be provided as a script in the TH PREFS file.

Overall datafile structure

Datafiles are "free format", with spacing and distribution of the information into lines being up to the user. Well laid out and annotated datafiles are easier to read and easier to debug. Annotation is easy: anything after a "%" on a line of input is *ignored* by THERMOCALC.

Datafiles are in three sections, separated by "∗"s. The first contains the information on the phases involved, the second contains scripts specific to this set of phases, and the third is a storage area not read by THERMOCALC. The information below pertains to the first section; script information is presented separately.

Simple coding

Coding for modes 2 and 3, and for phases involving only one end-member in mode 1, are simple. At its simplest, a datafile for mode 3, to calculate a PT projection for NASH might be

```
jd ab pa ky q H2O
```

in other words just a list of the end-members (using the THERMOCALC abbreviations—see separate file).

---

[1]Whereas it might have seemed more convenient for the user to have the a-x relationships "hard-wired" in THERMOCALC, obviating the pain of having to code them, in most phases the form of the a-x relationships are not prescribed in the fitting of the experimental data to make the thermodynamic dataset, and there is often not a consensus on how they should be formulated.

In such a list of end-member phases, implicitly the activity of each end-member is unity. In the context of mode 2 calculations, a particular end-member is rarely the sole constituent of a phase, so if an activity (less than 1) is needed it is included after the end-member name. So the RP13 datafile looks like

```
% RP13
mu      0.64      cel     0.041    pa      0.48
phl     0.072     ann     0.023    east    0.051    naph    0.017
clin    0.050     ames    0.027    daph    0.0085
gr      0.016     py    0.00092    alm     0.25
an      0.49      ab      0.71
cz      0.68
cc q H2O CO2
```

In this case, the activities were calculated from microprobe analyses using the software, AX. The activities of cc and q are unity. The activities of H_2O and CO_2 are handled separately from this part of the datafile, either via scripts, or via answering questions that THERMOCALC asks when it is run with this datafile.

In performing average PT calculations (mode 2), uncertainties on the activities are also required. If these are known, then they can be included, following the activity. Often they are not known, and THERMOCALC provides default uncertainties at run-time, over which the user has some control.

End-members that can be constructed from the end-members in the thermodynamic dataset, can have their thermodynamics constructed using "make" and "DQF" which are described below.

Coding for a phase involving more than one end-member

The datafile structure for each phase is

1. mineral name abbreviation

2. number of end-members, n

3. the variables and their starting guesses ($n - 1$ of them)

4. the end-member proportion "terms" (n of them)

5. a code—either *sf* or *ideal*—indicating whether the phase is non-ideal (symmetric formalism) or not (ie ideal).

6. if the code is *sf*, then the macroscopic interaction energies (W's)

7. the number of site fraction terms, s

8. the site fraction "terms" (s of them)

9. the ideal mixing activity expressions (n of them), in terms of the site fractions

So, for a simple example, staurolite

```
% ------------------------------------

st 2

  x(st) 0.89                          % starting guess, x = 0.89

                                      % end-member proportions
        p(mst)  1 1    1  1 -1  x     %  1 - x
        p(fst)  1 1    0  1  1  x     %      x

  sf    % symmetric formalism non-ideality

        w(st) -8 0 0                  % w(st) = -8 + 0(T) + 0(P)

  2     % no of site fractions

        x(Mg,oct)  1 1    1  1 -1  x     %  1 - x
        x(Fe,oct)  1 1    0  1  1  x     %      x

    mst     1    1     x(Mg,oct)  4   % a(mst,ideal) = x(Mg,oct)^4

    fst     1    1     x(Fe,oct)  4   % a(fst,ideal) = x(Fe,oct)^4

% ------------------------------------
```

The coding that follows `p(mst)`, `p(fst)`, `x(Mg,oct)` and `x(Fe,oct)` are the simplest examples of what I call a "term" in the next section. The last two lines are the ideal mixing activities, each in terms of the normalisation constant, followed by the number of site fraction terms involved, followed by a series of site fractions and their powers (just one site fraction term in these cases).

Terms

The new datafile construction is built around what I call a "term", θ, which has the general form

$$\theta = \sum_{i=1}^{n} \prod_{j=1}^{m_i} \left(c_{ij} + \sum_{k=1}^{t_{ij}} m_{ijk} v_{ijk} \right) \tag{1}$$

in which c_{ij} and m_{ijk} can be integer *or* rational (eg $\frac{2}{3}$). (For afficionados of the old coding, note that now there is no "power" at the innermost level (it is now applied to whole "term"s), in comparison to datafiles pre v3).

The coding of a "term" is as before (n, then m, then sets of info corresponding to what is in the brackets). This form can handle, for example,

1. x, coded as

```
    1   1      0   1   1   x
```

2. $1 - x$, coded as

```
    1   1      1   1  -1   x
```

3. $x + \frac{2}{3}y$, coded as

```
    1   1      0   2   1   x 2/3 y
```

4. $1 + x(1 - y)$, coded as

```
    2   1      1   0
        2      0   1   1   x      1   1  -1   y
```

5. $1 - y - \frac{2}{3}Q - x(1 - \frac{1}{3}y)$, coded as

```
    2   1      1   2  -1   y -2/3 Q
        2      0   1  -1   x      1   1  -1/3 y
```

The term is viewed as $\left(1 - y - \frac{2}{3}Q\right) + \left(-x(1 - \frac{1}{3}y)\right)$, so $n = 2$, $m_1 = 1$, $c_{11} = 1$, $t_{11} = 2$, $m_{111} = -1$, $v_{111} = y$, $m_{112} = -2/3$, $v_{112} = Q$, and $m_2 = 2$, $c_{21} = 0$, $t_{21} = 1$, $m_{211} = -1$, $v_{211} = x$, $c_{22} = 1$, $t_{22} = 1$, $m_{221} = -1/3$, $v_{221} = y$.

The variables, x and y, in these expressions correspond to $x(phasename)$ and $y(phasename)$ in the phase description.

The following biotite example is more complex, involving order-disorder of Fe and Mg between one M1 site and two M2 sites as outlined in Powell & Holland (1999). Al is considered to enter M1, not M2. Breaking the coding up into parts:

1. The choice of compositional variables is up to the user, and in more complex phases, there will be various more or less sensible choices. The order parameter here is just (a multiple of) the difference between the Fe/(Fe + Mg) on M2, and the bulk Fe/(Fe + Mg) (i.e. x)

```
bi 4     % formulated as in Powell and Holland (1999)

 x(bi) 0.4       % bulk Fe/(Fe + Mg)
 y(bi) 0.25      % x(Al,M1)
 N(bi) 0.14      % 3(x - x(Fe,M2))
```

2. The end-member proportions allow THERMOCALC to calculate the composition of the biotite from the calculated values of x, y and N. One way to visualise this is to draw the sites as boxes

for a particular composition of the mineral, as done here

From this $p_{\text{east}} = y$, $p_{\text{ann}} = x_{\text{Fe,M2}}$, $p_{\text{phl}} = x_{\text{MgM1}}$, and by difference the proportion of the ordered end-member is $p_{\text{obi}} = 1 - y - x_{\text{Fe,M2}} - x_{\text{MgM1}}$. To find the site fractions in terms of x, y and N, start by rearranging the definition of N

$$x_{\text{Fe,M2}} = x - \frac{N}{3}$$

and so, by difference,

$$x_{\text{Mg,M2}} = 1 - x + \frac{N}{3}$$

Now write x as

$$x = \frac{x_{\text{Fe,M1}} + 2x_{\text{Fe,M2}}}{x_{\text{Fe,M1}} + 2x_{\text{Fe,M2}} + x_{\text{Mg,M1}} + 2x_{\text{Mg,M2}}}$$

and, as $x_{\text{Fe,M2}} + x_{\text{Mg,M2}} = 1$ and $x_{\text{Fe,M1}} + x_{\text{Mg,M1}} = 1 - x_{\text{Al,M1}} = 1 - y$, then

$$x = \frac{x_{\text{Fe,M1}} + 2x_{\text{Fe,M2}}}{3 - y}$$

Substituting from above and simplifying gives

$$x_{\text{Fe,M1}} = x(1 - y) + \frac{2N}{3}$$

and

$$x_{\text{Mg,M1}} = (1 - x)(1 - y) - \frac{2N}{3}$$

So the coding of the proportions is

```
p(phl)   2 2    1  1 -1    x    1  1 -1   y
         1      0  1 -2/3  N

p(ann)   1 1    0  2  1    x   -1/3  N

p(east)  1 1    0  1  1    y

p(obi)   2 2    0  1 -1    x    0  1  1  y
         1      0  1  1    N
```

The p's must be in the same order as in the ideal mixing activity expression below.

3. Next comes the treatment of non-ideality:

```
sf

W(phl,ann)    9  0  0
W(phl,east)  10  0  0
W(phl,obi)    3  0  0
W(ann,east)  -1  0  0
W(ann,obi)    6  0  0
W(east,obi)  10  0  0
```

The W's have an order that is predicated on the order of the end-members used for the ideal mixing activities (so if one numbers the end-members, 1, 2, 3 etc, then the W's are in order 12, 13, ..., 23, 24, ...)

4. The site fraction expressions are the ones derived above:

```
7      % no of site fractions

x(Al,M1)  1 1    0  1   1    y

x(Fe,M1)  2 2    0  1   1    x    1  1 -1  y
                 1  0   1 2/3  N

x(Mg,M1)  2 2    1  1  -1    x    1  1 -1  y
                 1  0   1 -2/3 N

x(Fe,M2)  1 1    0  2   1    x -1/3   N

x(Mg,M2)  1 1    1  2  -1    x  1/3   N

x(Al,T1)  1 1    1/2 1  1/2  y

x(Si,T1)  1 1    1/2 1 -1/2  y
```

5. The ideal mixing activities are then formulated in terms of the site fractions:

```
phl     4  4    x(Mg,M1) 1   x(Mg,M2) 2   x(Al,T1) 1   x(Si,T1) 1
        check 0 0 0

ann     4  4    x(Fe,M1) 1   x(Fe,M2) 2   x(Al,T1) 1   x(Si,T1) 1
        check 1 0 0

east    1  3    x(Al,M1) 1   x(Mg,M2) 2   x(Al,T1) 2
        check 0 1 0

obi     4  4    x(Fe,M1) 1   x(Mg,M2) 2   x(Al,T1) 1   x(Si,T1) 1
        check 1/3 0 1
        make       2    phl 2/3 ann 1/3
        DQF      -10.73  0   0
```

The "check", "make" and "DQF" info is described below.

Other coding

The other coding (DQF, make and check) are unchanged from before (see the web description), except: There is an important change to "make": a divider (the reaction coefficient of the end-member being "made") is no longer provided. Instead the reaction coefficients of the (other) end-members involved can be rational, as required (see biotite example above).

Consistency checking

Consistency checking of the ideal mixing activity and bulk formulations are done if the values of the variables for the pure end-member are coded after the 'check' key-word in the datafile. So, in the above example, for obi

```
check 1/3 0 1
```

means that, in end-member obi, $x = \frac{1}{3}$, $y = 0$ and $N = 1$.

Making end-members

Only an independent set of end-members for each mineral group is included in the dataset. However sometimes the thermodynamic properties of a dependent end-member are required. In the simple case, in which the phase is an ideal mixture of the end-members, the Gibbs energies of the dependent end-member, and those in the independent set used to make it, are coplanar, and the dependent end-member is completely specified in this way. The situation is more complex if the phase is non-ideal because the Gibbs energy of the dependent end-member must be augmented with a linear combination of the interaction energies (see Powell & Holland, 1999). In the simplest case, the information provided, following the keyword 'make', is just the end-member names and their reaction coefficients for the reaction that makes the dependent end-member.

Another use of 'make' is shown in the biotite example. the ordered end-member, obi, is made from phlogopite and annite, then its thermodynamic properties adjusted using DQF (see next). A more brutal example of this use is to make a ferric opx end-member (mots, $MgFe^{3+}AlSiO_6$), with

```
make       3       mgts 1    cor -1/2    hem 1/2
DQF     22  0  0
```

Note that the 3 following "make" refers to the number of end-members in the "equation" to make mots.

DQF

When using "make", or to implement non-ideality via Darken's Quadratic Formalism (DQF) approach (Powell, 1987; Holland and Powell, 1992), it is appropriate to adjust the Gibbs energy of an end-member The inclusion of such information is flagged with the key-word, 'DQF'. It involves 3 numbers, a, b and c, which, in the form, $a + bT + cP$, is added to the Gibbs energy of the end-member.