Part 5. Multisite mixing models: built-in functions in GEM-Selektor v.3.4 code

Treatment of condensed symmetric non-electrolyte solutions in GEM-Selektor code is based on the default expression for the chemical potential of a j-th end member:

$$\upsilon_j = \frac{g_{j,T,P}}{RT} + \ln x_j + \ln f_j \tag{5.1-1},$$

where $g_{j,T,P}$ is the molar apparent Gibbs energy function at T,P of interest; $x_j = {n_j^{(x)}}\big/{n_k^{(x)}}$ is the end member mole fraction (n_j is the primal mole quantity of j-th species and $n_k = \sum x_j$, $j \in L_k$ is the total mole quantity of the k-th phase), and f_j is a 'real' activity coefficient.

If all $f_j = 1$ at any composition then eqn (5.1-1) reduces to the Raoult's law of simple ideal mixing, where the activity of j-th end-member equals its mole fraction. This definition of ideal mixing, corresponding to the configurational entropy of random mixing on a single type of structural sites, is rather narrow. In a broader sense, ideal solutions can be defined as having zero excess entropy, zero enthalpy and zero volume of mixing (Oates, 1966; Price, 1985; Anderson and Crerar, 1993). Microscopic moieties (usually ions or atoms) can substitute each other independently on one or several types of structural sites, often called 'sublattices' (Hillert, 1998; Lucas et al., 2007). Sublattices formally represent long-range ordering (LRO).

Eq (5.1-1) corresponds to the simplest case where two or more moieties substitute each other on one site type, and there is one moiety per formula of end member. This is the case e.g. for carbonate $(A,B,...)CO_3$ or sulfate $(A,B,...)SO_4$ solid solutions, but it is clearly insufficient for most of rock-forming silicate or oxide minerals where mixing occurs on several sublattices simultaneously, and can be complicated by short-range ordering (SRO) effects. In such ideal 'multisite mixing', the activity a_j of j-th end member is not simply its mole fraction x_j, but a more complex function of 'site fractions' $y_{s,m}$ of m-th substituting moiety on s-th site type in k-th solution phase. This corresponds to a more complex form of the configurational entropy, and causes the 'reciprocal' relations between end members and their standard-state properties.

Here we need to introduce a set Ξ_k of indexes of sublattices (structural site types) in the k-th phase, where $n(\Xi_k)$ is the number of site types indexed with $s \in \Xi_k$. Further, we introduce a set M_s of indexes of $n(M_k)$ moieties that can appear on s-th site in k-th phase structure; moieties will be indexed with $m \in M_k$. Let $\eta_{j,s,m}$ be a 'moiety-site multiplicity', i.e. the number of moles of m-th moiety on s-th site in one mole of j-th end member at standard state ($j \in L_k$), and η_s be the number of s-th sites in one formula unit of any end member of the k-th phase ($s \in \Xi_k$) ('site multiplicity'). Note that the same species (e.g. ion M+), if it appears on two different sublattices, must be considered as two different moieties; the same applies to vacancies.

5.1. *Ideal part of the end-member activity term*

Now, the approximation of chemical potential of j-th end member of k-th phase takes the form

$$\upsilon_j = \frac{g_{j,T,P}}{RT} + \ln a_j^{(con)} + \ln f_j \qquad j \in L_k \tag{5.1-2},$$

where f_j is the 'macroscopic' end member activity coefficient, and $a_j^{(con)}$ is the ideal part of activity related to the partial molar configurational entropy:

$$-R\ln a_j^{(con)} = \Delta\overline{S}_j^{(con)} - \Delta\overline{S}_j^{(con),o} \tag{5.1-3}.$$

Following the derivation in (Price, 1985),

$$\Delta \bar{S}_j^{(con)} = -R \sum_s \sum_m \eta_{j,s,m} \ln y_{s,m} \qquad m \in M_k \quad s \in \Xi_k \quad j \in L_k \qquad (5.1\text{-}4),$$

where the site fraction of m-th moiety on s-th site in k-th phase is

$$y_{s,m} = \frac{n_{s,m}}{\sum_m n_{s,m}} \qquad m \in M_k \qquad (5.1\text{-}5).$$

here $n_{s,m}$ is the number of moles of m-th moiety on s-th sublattice sites in the 'real' state. At standard state, the s-th site fraction of m-th moiety in j-th end member is

$$y_{j,s,m}^o = \frac{\eta_{j,s,m}}{\eta_s} \qquad (\text{'site multiplicity'} \ \eta_s = \sum_m \eta_{j,s,m}) \qquad (5.1\text{-}6),$$

and the end member has a configurational entropy term itself:

$$\Delta \bar{S}_j^{(con),o} = -R \sum_s \sum_m \eta_{j,s,m} \ln y_{j,s,m}^o \qquad (5.1\text{-}7)$$

Combination of eqs (5.1-4, 5.1-7 and 5.1-3) yields

$$\ln a_j^{(con)} = \sum_s \sum_m \eta_{j,s,m} \cdot \ln \left(\frac{y_{s,m}}{y_{j,s,m}^o} \right) \qquad m \in M_k \quad s \in \Xi_k \quad j \in L_k \qquad (5.1\text{-}8)$$

The site fraction $y_{s,m}$ (eq 5.1-5) in the k-th phase can be expressed through mole fractions of end members x_j because

$$n_{s,m} = \sum_{j_k} \eta_{j_k,s,m} n_{j_k} = n_k \sum_{j_k} \eta_{j_k,s,m} x_{j_k} \qquad j_k \in L_k \qquad (5.1\text{-}9),$$

where n_{jk} is the primal mole amount of j_k-th end member, and $n_k = \sum_{j_k} n_{j_k}$ (here, summations run over all end members in k-th phase). Using $\eta_s = \sum_m \eta_{j,s,m}$, eq (5.1-5) takes the form:

$$y_{s,m} = \frac{\sum_{j_k} \eta_{j_k,s,m} x_{j_k}}{\sum_m \sum_{j_k} \eta_{j_k,s,m} x_{j_k}} = \frac{1}{\eta_s} \sum_{j_k} \eta_{j_k,s,m} x_{j_k} \qquad j_k \in L_k \quad m \in M_k \quad s \in \Xi_k \qquad (5.1\text{-}10).$$

Substituting this and eq (5.1-6) into eq (5.1-8) results in the ideal activity of j-th end member

$$\ln a_j^{(con)} = \sum_s \sum_m \eta_{j,s,m} \{ \ln(\sum_{j_k} \eta_{j_k,s,m} x_{j_k}) - \ln \eta_{j,s,m} \} \qquad j_k \in L_k \qquad (5.1\text{-}11).$$

Now, eq (5.1-2) for the end member chemical potential takes a more general form:

$$\upsilon_j = \frac{g_{j,T,P}}{RT} + \sum_s \sum_m \eta_{j,s,m} \{ \ln(\sum_{j_k} \eta_{j_k,s,m} x_{j_k}) - \ln \eta_{j,s,m} \} + \ln f_j \qquad (5.1\text{-}12).$$

At $n(\Xi) = 1$ and all $\eta_{j,s,m}$ numbers equal to unity, eq (5.1-11) reduces to $\ln x_j$ – the simple ideal single-site mixing activity contribution - and eq (5.1-12) reduces to eq (5.1-1). In principle, any multi-site mixing model can be put on top of the $\ln x_j$ term in eq (5.1-1) by introducing a 'configurational entropy' term λ_j, also called 'fictive activity coefficient':

$$\upsilon_j = \frac{g_{j,T,P}}{RT} + \ln x_j + \ln \lambda_j + \ln f_j \qquad (5.1\text{-}13)$$

where

$$\ln \lambda_j = \sum_s \sum_m \eta_{j,s,m} \{ \ln(\sum_{j_k} \eta_{j_k,s,m} x_{j_k}) - \ln \eta_{j,s,m} \} - \ln x_j \qquad (5.1\text{-}14).$$

In earlier versions of GEMS, this could be done only in Phase definition scripts which contained all the specific $\eta_{j,s,m}$ numbers for a particular sublattice mixing model. This slowed calculations down and precluded such phase definitions from being used in GEMS3K kernel and, thus, in any reactive mass transport or other coupled codes. This work represents an extension that enables built-in ideal or non-ideal multi-site mixing models also in GEMS3K. For this extension in TSolMod class, it is necessary to use an array of $\eta_{j,s,m}$ numbers arranged as a sequence of $n(L_k)$ matrices $\left\| \eta_{s,m} \right\|_j$, each having $n(\Xi_k)$ rows and $n(M_k)$ columns. These matrices will be arranged in (part of) a MoiSN array (to be created in the MULTI data structure) then passed as a pointer into TSolMod class instance. The moieties will be arranged in a common list (without duplicate entries) for the solution phase. If the same cation or anion occurs in different sublattices then it must be considered as a different moiety on each sublattice site. For example, in ferrites, divalent and trivalent cations may be redistributed between tetrahedral and octahedral sites, e.g. in Ni-ferrite: $(Ni_{1-z}Fe^{III}_z)_{Tet}(Ni_zFe^{III}_{2-z})_{Oct}O_4$, where $0 < z < 1$ is the extent of inversion which depends on temperature. At given temperature, e.g. setting $z = 0.8$, the Fe-Ni ferrite system may be represented using two end members: $(Ni_{0.2}Fe^{III}_{0.8})_{Tet}$ $(Ni_{0.8}Fe^{III}_{1.2})_{Oct}O_4$ and $(Fe^{II}_{0.2}Fe^{III}_{0.8})_{Tet}$ $(Fe^{II}_{0.8}Fe^{III}_{1.2})_{Oct}O_4$. Note that Ni cation on the Tet sublattice is not the same moiety as Ni cation on the Oct sublattice. If vacancies are considered on different sites, the same applies to vacancies.

5.1.1. *Coding sublattices in end-member formulae*

The most versatile way of providing the information about moieties occupancies on sites in the Phase definition seems to be through the formulae of end members. To do this, the formulae syntax needs to be extended. For example, the usual formulae of feldspar end members are

`KAlSi3O8` (microcline); `NaAlSi3O8` (albite); `CaAl2Si2O8` (anorthite) etc.

This format does not contain information about sites and moieties that substitute on them. However, this information can be incorporated e.g. in the following way:

`{K}:{Al}{Si}3:O8`   `{Na}:{Al}{Si}3:O8`   `{Ca}:{Al}2{Si}2:O8`

Essential feature here is that each moiety (incl. vacancy `Va`) is taken in the {} braces; the colon ':' separates sublattice sites which must follow the same order in all end members. The $\eta_{j,s,m}$ numbers appear just as stoichiometry coefficients after moieties. Upon parsing the formula, the number of sublattices is counted by encountered colons ':' (two in the above example), and the moieties are collected in a list along with their site-moiety multiplicity η numbers and the site index. This list can then be merged into the overall list of $n(M_k)$ moieties for the phase, and the η numbers then assigned to $\left\| \eta_{s,m} \right\|_j$ tables, the sequence of which forms the MoiSN array. The same ion can be represented as different moieties by adding the moiety index after species name. In the above example of ferrite with cation inversion, the formulae of end members will look like this (default valence of IC Ni is 2 and that of IC Fe is 3):

`{Ni}0.2{Fe}0.8:{Ni}0.8{Fe}1.2:O4`

`{Fe|2|}0.2{Fe|3|}0.8:{Fe|2|}0.8{Fe|3|}1.2:O4`

In the former end member, `Ni` is present on both sites, so the program creates two moieties: `[Ni]0` and `[Ni]1`; it also creates two moieties for divalent Fe: `[Fe]0` and `[Fe]1`. By analyzing the latter end member, `[Fe|3|]0` and `[Fe|3|]1` will be added to list of moieties, and so on.

Chemical formulae of DCs are parsed in GEM-Selektor using the syntax that is now extended, as shown below in boldface:

```
<formula> ::= <fterm> | <fterm> <charge>
<fterm>   ::= <icterm> | <fterm> <icterm> |
              <site_term> : | <fterm> <site_term> :
<site_term> ::= <moiety> | <site_term> <moiety>
<moiety>  ::= {<elem>} | {<elem>} <elem_st_coef>
<icterm>  ::= <elem> | <elem> <elem_st_coef>
<elem>    ::= (<fterm>) | [<fterm>] |
              <isotope_mass><icsymb><valence> |
              <isotope_mass><icsymb> |
              <icsymb><valence> | <icsymb>
<charge>  ::= +<real> | -<real> | + | - | @
<elem_st_coef> ::= <real>
<icsymb>  ::= <Capital_letter> | <icsymb><lcase_letter>
              | <icsymb>_ | Va
<valence> ::= |-<integer>| | |+<integer>| | |<integer>|
<isotope_mass> ::= /<integer>/
<integer> ::= <num>
<num>     ::= <digit> | <num><digit>
<digit>   ::= 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9
<real>    ::= <num>.<num> | <num>. | .<num> | <num>
```

In the **`<elem>`** term, up to three levels of nesting is allowed. The vertical bar "|" means "or", i.e. one selection from several possible values or terms. The "`Va`" in <icsymb> means "vacancy" and should be ignored everywhere except where moieties and site terms are analyzed. When the formula parser is called for checking the DComp or ReacDC formulae, all the { and } brackets must be replaced with (and) respectively, and colons : removed.

Otherwise, the site index should be incremented upon encountering the colon ':'. The newly encountered moieties (including `Va`) is put into square brackets, appended with the sublattice index, and then added to the list of moieties, if not already present there, and the moiety index and site index are then used for inserting the respective occupancy number (if given after '}', 1 by default) into the occupancy table, which must exist for all end members in the phase. The strict order of site terms is essential. The total of moiety occupancies (the site multiplicity) for a given site is checked to be the same for this site term in all end members (see eq 5.1-6).

5.2. *Non-ideal site-interaction part of the end-member activity term*

To represent intra-site non-ideal interactions for substitutions of moieties on sublattices, the approach of Berman and Brown (1984), Berman (1990) is applied because of its relative simplicity and compatibility with 'petrologic' style of defining end members with more than one moiety on a sublattice site. So far, this approach ignores cross-site interactions (which at least partially can be accounted for using the reciprocal terms).

In our notation, the ideal contribution to the end member activity used in Berman's models,

$$a_j^{(con)} = \prod_s \prod_m \left(y_{s,m} \frac{\eta_s}{\eta_{j,s,m}} \right)^{\eta_{j,s,m}} \qquad m \in \mathrm{M}_k \quad s \in \Xi_k \quad j \in L_k \qquad (5.2\text{-}1)$$

is shown to be equivalent to eq (5.1-8) because from eq (5.1-6), $\dfrac{1}{y_{j,s,m}^o} = \dfrac{\eta_s}{\eta_{j,s,m}}$.

The non-ideal contribution to end-member activity (see eq 5.1-2) is defined as the sum of

$$RT \ln f_j = \sum_s f_{j,s} \qquad s \in \Xi_k \quad j \in L_k \qquad (5.2\text{-}2)$$

Furthermore,

$$f_{j,s} = RT \sum_m \ln \gamma_{j,s,m}$$

$$RT \ln \gamma_{j,s,m} = y_{j,s,m}^o \sum_t y_d y_e y_f W_{def,s} \left(Q_{s,m} \frac{y_{j,s,m}^o}{y_{s,m}} - \Theta \right) \qquad t \in \mathrm{W}_s \qquad (5.2\text{-}3)$$

where the index *m* refers only to moieties that are present on *s*-th site in *j*-th end member; $Q_{s,m}$ is the number of $_{d,\ e,\ f}$ subscripts equal to *m* (0, 1 or 2); Θ is the order of interaction (symmetric: 1; asymmetric: 2); and W_s is the subset of indexes *p* of entries in the list of interaction parameters that refer to *s*-th sublattice. The $_{d,\ e,}$ or $_f$ subscript can be an index *m* of any moiety that exists on the *s*-th site. Hence, $W_{def,s}$ is the interaction parameter for a particular combination of moieties substituting each other on the *s*-th site. The *T,P* dependence of the interaction parameter is represented, as usual, by three coefficients *a, b,* and *c* (J,K,bar):

$$W = a - b{\cdot}T + c{\cdot}P \qquad (5.2\text{-}4)$$

In case of symmetric interaction, the $_f$ subscript does not exist, and eq (5.2-3) takes the form

$$RT \ln \gamma_{j,s,m} = y_{j,s,m}^o \sum_t y_d y_e W_{de,s} \left(Q_{s,m} \frac{y_{j,s,m}^o}{y_{s,m}} - 1 \right) \qquad (5.2\text{-}5)$$

where $y_{j,s,m}^o = \dfrac{\eta_{j,s,m}}{\eta_s}$ (eq 5.1-6). For the asymmetric interactions,

$$RT \ln \gamma_{j,s,m} = y_{j,s,m}^o \sum_t y_d y_e y_f W_{def,s} \left(Q_{s,m} \frac{y_{j,s,m}^o}{y_{s,m}} - 2 \right) \qquad (5.2\text{-}6)$$

where either d and e or e and f indexes refer to the same moiety on s-th site in the case of a pseudo-binary interaction, and all three indexes are different in a ternary interaction.

In addition to tables $\left\|\eta_{s,m}\right\|_j$ and $\left\|y_{s,m}\right\|$ taken over from the calculation of ideal contributions to end member activities (see Section 5.1), another set of tables $\left\|f_{j,s}\right\|_j$ is needed in the algorithm to collect the contributions to terms $RT\ln\gamma_{j,s,m}$ (see eq 5.2-3).

In the TSolMod class implementation, the setup of interaction parameters (IP) is done in the Phase definition record. Upon calculation of equilibria, the IP list is passed automatically into the TSolMod instance for calculation of end-member activity coefficients. In the microscopic non-ideality model, the IP list needs to be arranged in the following format:

IP index   Site index Moiety indexes   IP coefficients (J, K, bar)

p	s	d	e	f	a	b	c
0	0	0	1	-1	-6000	-1.3	0.4
1	1	3	3	4	17000	5.1	0.1
…	...	...	...	...	...	...	...

IP coefficients refer to eq (5.2-4), $W_{def,s} = a - b{\cdot}T + c{\cdot}P$ (note the minus sign at bT term), T in K, P in bar, a in J/mol.

Example: Muscovite (white mica) solid solution (Parra et al. 2002)

End-members:
muscovite (mu), celadonite (cel), ferrous celadonite (fcel), Na-celadonite (ncel), Na-ferrous celadonite (nfcel), paragonite (pa), and pyrophyllite (prl).

Moiety- sublattice site allocation table (encoded in end-member formulae).

Sublattice / End member	A (0)	M23 (1)	T2 (2)	T1 (3)	End-member formula (as written in DComp record)
(0) fcel	K	Al Fe	Si_2	Si_2	{K}:{Al}{Fe}:{Si}2:{Si}2:O10(OH)2
(1) cel	K	Al Mg	Si_2	Si_2	{K}:{Al}{Mg}:{Si}2:{Si}2:O10(OH)2
(2) mu	K	Al_2	Al Si	Si_2	{K}:{Al}2:{Al}{Si}:{Si}2:O10(OH)2
(3) nfcel	Na	Al Fe	Si_2	Si_2	{Na}:{Al}{Fe}:{Si}2:{Si}2:O10(OH)2
(4) ncel	Na	Al Mg	Si_2	Si_2	{Na}:{Al}{Mg}:{Si}2:{Si}2:O10(OH)2
(5) pa	Na	Al_2	Al Si	Si_2	{Na}:{Al}2:{Al}{Si}:{Si}2:O10(OH)2
(6) prl	Va	Al_2	Si_2	Si_2	{Va}:{Al}2:{Si}2:{Si}2:O10(OH)2

Sublattice site interaction parameters (J, K, bar).

$W_{KKNa,A} = 19460 - 1.65 \cdot T - 0.46 \cdot P$

$W_{KNaNa,A} = 12230 - 5.0 \cdot T + 0.67 \cdot P$

$W_{KKVa,A} = 35000 + 25.0 \cdot T - 0.85 \cdot P$

$W_{KVaVa,A} = 45000 + 10.0 \cdot T - 0.85 \cdot P$

$W_{NaNaVa,A} = 40000 + 5.0 \cdot T$

$W_{NaVaVa,A} = 40000 + 5.0 \cdot T$

$W_{KNaVa,A} = 95480 + 19.13 \cdot T - 0.745 \cdot P$

$W_{AlMg,M23} = -30500 + 15.0 \cdot T + 0.78 \cdot P$

$W_{AlFe,M23} = -5500 + 15.0 \cdot T + 0.65 \cdot P$

$W_{MgFe,M23} = 0$

$W_{AlSi,T2} = 0$

In GEMS Phase window (Page 3), the recalculation results in the following allocation table:

End member- moiety multiplicity number table (multiplicity numbers can be non-integer).

m	0	1	2	3	4	5	6	7	8
j	{K}0	{Al}1	{Fe]}1	{Si}2	{Si}3	{Mg}1	{Al}2	{Na}0	{Va}0
0	1	1	1	2	2	0	0	0	0
1	1	1	0	2	2	1	0	0	0
2	1	2	0	1	2	0	1	0	0
3	0	1	1	2	2	0	0	1	0
4	0	1	0	2	2	1	0	1	0
5	0	2	0	1	2	0	1	1	0
6	0	2	0	2	2	0	0	0	1

Indexes m and j are shown for the reader's convenience

To set up interaction parameters, remake the Phase record and set the Berman (B) built-in (S) model with total number of input interaction parameters (in this example, 11). In Page 3 of Phase window, you will see a combined list filled with zeros.

As shown below, for each interaction parameter, the indexes of sublattice s and moieties d,e,f should be entered into four ipxT colums (for a symmetric interaction parameter, index f should be set to -1). For each moiety, its sublattice index can be found right after its name, e.g. {K}0. For example, the parameter $W_{KKNa,A} = 19460 - 1.65 \cdot T - 0.46 \cdot P$ will be set as shown below:

s	d	e	f	a	b	c
ipxT[0]	ipxT[1]	ipxT[2]	ipxT[3]	ph_cf[0]	ph_cf[1]	ph_cf[2]
0	0	0	7	19460	-1.65	-0.46
...	...	...	...	...	...	...

The complete list for this example is shown below. The interaction parameters (rows) can be given in an arbitrary sequence. Two rows with zero IPs can, in principle, be dropped (by remaking the Phase record and setting the number of rows to 9 instead of 11). However, the sequence of end members and moieties is essential; if it changes, the indexation in ipxT columns must be revised according to the new moiety allocation table.

Interaction parameters setup in Phase record for the Muscovite example.

ipxT[0]*	ipxT[1]	ipxT[2]	ipxT[3]	ph_cf[0]	ph_cf[1]	ph_cf[2]
0	0	0	7	19460	-1.65	-0.46
0	0	7	7	12230	-5	0.67
0	0	0	8	35000	25	-0.85
0	0	8	8	45000	10	-0.85
0	7	7	8	40000	5	0
0	7	8	8	40000	5	0
0	0	7	8	95840	19.13	-0.745
1	1	5	-1	-30500	15	0.78
1	1	2	-1	-5500	15	0.65
1	5	2	-1	0	0	0
2	6	3	-1	0	0	0

*This column contains the index of sublattice

Worked test examples of multi-site solid solution models with Berman microscopic non-ideality will be provided in the BermanMSS test project and its documentation (TBD).

5.3. Reciprocal part of the end-member activity term

The typical multi-site solid solution with two or more sublattices and two or more moieties substituted on each sublattice is called *reciprocal*. Meaning of this term is best illustrated using the system {A,B}{X,Y} (e.g. Wood and Nicholls, 1978; Hillert, 2001). All possible compositions can be represented on a 'composition square' formed by four possible end members: AX, BX, AY, BY. There are four end members, but only three are needed to describe any bulk composition of this solid solution phase. For instance, the center of the square can be obtained by mixing equal amounts of either AX and BY or AY and BX, hence the name 'reciprocal'. Any three of four end members can be declared as 'independent', while the remaining one will be 'dependent' because it's composition and thermodynamic properties can be obtained from that of 'independent' end members and that of a *reciprocal reaction*

$$AY + BX \Leftrightarrow AX + BY \tag{5.3-1}$$

with a standard molar Gibbs energy effect of reaction

$$\Delta^o G_{(A,B)(D,E)} = G^o_{AX} + G^o_{BY} - G^o_{AY} - G^o_{BX}$$

For example, if the BE end member is declared as 'dependent' then it's standard molar Gibbs energy can be obtained from eq (5.3-1) as $G^o_{BY} = \Delta^o G_{(A,B)(X,Y)} + G^o_{AY} + G^o_{BX} - G^o_{AX}$, provided that the Gibbs energy effect of the reciprocal reaction is known (e.g. fitted from the experimental data). There is also a condition of internal equilibrium: $\mu_{AX} + \mu_{BY} = \mu_{AY} + \mu_{BX}$ (Wood and Nicholls, 1978). Following the derivations in this paper, assuming random ideal mixing on both sublattices, the chemical potentials of end members in this system are:

$$\mu_{AX} = G^o_{AX} + RT\ln\left(y_{0,A}y_{1,X}\right) - y_{0,B}y_{1,Y}\Delta G^o_{(A,B)(X,Y)}$$

$$\mu_{BY} = G^o_{BY} + RT\ln\left(y_{0,B}y_{1,Y}\right) - y_{0,A}y_{1,X}\Delta G^o_{(A,B)(X,Y)}$$

$$\mu_{AY} = G^o_{AY} + RT\ln\left(y_{0,A}y_{1,Y}\right) + y_{0,B}y_{1,X}\Delta G^o_{(A,B)(X,Y)}$$

$$\mu_{BX} = G^o_{BX} + RT\ln\left(y_{0,B}y_{1,X}\right) + y_{0,A}y_{1,Y}\Delta G^o_{(A,B)(X,Y)}$$

Comparing these with the general form of the chemical potential, we see that the terms like $y_{0,B}y_{1,Y}\Delta G^o_{(A,B)(X,Y)}$ stay in place of one more activity coefficient like γ_{AX}:

$$RT\ln\gamma_{AX} = -y_{0,B}y_{1,Y}\Delta G^o_{(A,B)(X,Y)}$$

$$RT\ln\gamma_{BY} = -y_{0,A}y_{1,X}\Delta G^o_{(A,B)(X,Y)}$$

$$RT\ln\gamma_{AY} = +y_{0,B}y_{1,X}\Delta G^o_{(A,B)(X,Y)}$$

$$RT\ln\gamma_{BX} = +y_{0,A}y_{1,Y}\Delta G^o_{(A,B)(X,Y)}$$

Thus, the complete form of the end-member chemical potential will be

$$\mu_j = \mu_j^o + RT \ln x_j + RT\ln\lambda_j + RT\ln f_j + RT\ln\gamma_j \qquad (5.3\text{-}2)$$

where $RT\ln\gamma_j$ is the reciprocal activity coefficient term, which turns to zero (ideal behaviour) only if the Gibbs energy effect of the reciprocal reaction (5.3-1) $\Delta G^o_{(A,B)(X,Y)} = 0$. This formalism can be extended for the case when several moieties substitute on each of two sublattices (Wood and Nicholls, 1978; Aranovich, 1991). The number of end members in this case is M_0M_1; the number of independent end members is M_0+M_1-1; and the number of all possible reciprocal reactions is the number of combinations $C(M_0M_1, M_0+M_1-1)/4$.

Unfortunately, it is very difficult to generalize this "reciprocal reaction" formalism to solid solutions with simultaneous substitutions in three or more sublattices. For the general case, Sundman and Ågren (1981) suggested a method related to the site fractions matrix Y made of the $y_{s,m}$ elements (see eq 5.1-5). The total Gibbs energy per mole of k-th phase is represented as

$$G_k = G_k^{ref} - TS_{k,id}^{conf} + G_k^{EX} \qquad (5.3\text{-}3)$$

where $G_k = \sum_j \mu_j x_j$; $S_{k,id}^{conf} = \sum_j x_j\Delta\bar{S}_j^{(con)}$ where $\Delta\bar{S}_j^{(con)}$ is given by eq (5.1-4); $G_k^{EX} = RT\sum_j x_j\ln f_j$ where $\ln f_j$ is given by eq (5.2-2); and G_k^{ref} is the so-called frame-of-reference Gibbs energy (Sundman and Ågren, 1981; Hillert, 1998) .

For the substitution on one sublattice, $G_k^{ref} = \sum_{m_0}^{M_0} y_{0,m_0} G_{m_0}^o = \sum_j x_j G_j^o$.

For simultaneous substitutions on two sublattices, $G_k^{ref} = \sum_{m_0}^{M_0}\sum_{m_1}^{M_1} y_{0,m_0}y_{1,m_1} G_{m_0m_1}^o$ where $G_{m_0m_1}^o$ is the standard-state Gibbs energy of the end-member compound made of a moiety m_0 on sublattice 0 and moiety m_1 on sublattice 1.

For three sublattices, $G_k^{ref} = \sum_{m_0}^{M_0}\sum_{m_1}^{M_1}\sum_{m_2}^{M_2} y_{0,m_0}y_{1,m_1}y_{2,m_2} G_{m_0m_1m_2}^o$, and so on.

Sundman and Ågren (1981) suggested the following general notation:

$$G_k^{ref} = \sum_I G_I^o \wp_I(Y) \qquad (5.3\text{-}4)$$

where I is the "component array" that defines one moiety on each sublattice, and $\wp_I(Y)$ stands for the corresponding product of site fractions from the Y matrix. The CEF (compound energy formalism) eq (5.3-3) assumes that in any end member, there is exactly one moiety on each sublattice. In the formalism considered in Sections 5.1 and 5.2, end members with fractional site occupation are allowed, i.e with two or more moieties on one sublattice with end member site fractions $y_{j,s,m}^o = \frac{\eta_{j,s,m}}{\eta_s}$ (see eq 5.1-6). To account for this case, we assume that $\wp_{I0}(Y) = \wp_j(Y)$ is the product of sums $y_{s,\sigma}^{(j)}$:

$$y_{s,\sigma}^{(j)} = \sum_m y_{j,s,m}^o y_{s,m} = \frac{1}{\eta_s}\sum_m \eta_{j,s,m}y_{s,m}. \qquad (5.3\text{-}5)$$

Clearly, the CEF simply means that there is only one m-th non-zero element $\eta_{j,s,m} = \eta_s$ per j-th end member in s-th sublattice, so the $y_{s,\sigma}^{(j)}$ turns into y_{s,m_j}, the site fraction of moiety m_j which is present on s-th site in j-th end member formula.

The chemical potential of j-th end member can be obtained as a partial derivative of eq (5.3-3) on the amount of j-th end member,

$$\mu_j = G_k + \sum_s \left(\frac{\partial G_k}{\partial y_{s,m_j}} - \sum_m \frac{\partial G_k}{\partial y_{s,m}} \right)_{T,P,n_{l \neq j}} \tag{5.3-6}.$$

(Sundman and Ågren, 1981). Because of the additivity of G_k function, one can write

$$RT\ln\gamma_j = G_k^{ref} + \sum_s \left(\frac{\partial G_k^{ref}}{\partial y_{s,m_j}} - \sum_m y_{s,m} \frac{\partial G_k^{ref}}{\partial y_{s,m}} \right) - G_j^o, \quad s \in \Xi_k, \quad m \in M_k \tag{5.3-7}$$

where

$$\frac{\partial G_k^{ref}}{\partial y_{s,m_j}} = \sum_l \wp_l(Y) \cdot \frac{\delta_{j,s,m_j}}{y_{s,m}} \cdot G_l^o, \quad l \in L_k \tag{5.3-8}$$

$$\frac{\partial G_k^{ref}}{\partial y_{s,m}} = \sum_l \wp_l(Y) \cdot \frac{\delta_{l,s,m}}{y_{s,m}} \cdot G_l^o, \quad l \in L_k \tag{5.3-9}$$

Here, $\delta_{j,s,m}$ is the special form of Kronecker's delta: $\delta_{j,s,m} = 1$ if $\eta_{j,s,m} \neq 0$, else $\delta_{j,s,m} = 0$.

There are two limitations of the formalism eqs (5.3-4) to (5.3-9): (i) partially-substituted end members are not allowed; (ii) all possible end members must be included into the model. If these two conditions are not satisfied then the reciprocal energy contribution terms cannot be computed correctly and must be disabled.

From eq. (5.3-3) it is evident that the configurational entropy contribution is the same in Berman approach and in CEF model, whereas the excess energy contribution in Berman model is not the same as in CEF. Also, the reciprocal energy contributions were ignored in most known applications of Berman's model (maybe except those for systems with two sublattices, cf. Aranovich 1991). Therefore, in TSolMod library, two models of this kind are implemented: Berman (coded 'B') and CEFmod (coded '$'), see Appendix 5.1.

5.4. *Non-ideal excess energy part of the end-member activity term in CEF model*

In the Compound Energy Formalism (CEF), cf.(Sundman and Ågren 1981; Hillert 1988, 2001), the total Gibbs energy G_k per mole of *k*-th phase is represented by eq (5.3-3)

$$G_k = G_{k,ref} - TS_{k,id}^{conf} + G_{k,ex}, \quad \text{with the } G_{k,ref} \text{ term given by eq (5.3-4)};$$

$$S_{k,id}^{conf} = \sum_j x_j \bar{S}_j^{(conf)} \quad \text{where } \bar{S}_j^{(conf)} \text{ is given by eq (5.1-4)},$$

in which $\eta_{j,s,m} = \eta_s$ for any end member where exactly one moiety on each sublattice must exist. Both terms are the same as considered above for Berman's approach, but the excess Gibbs energy term is represented in a different way than eq (5.2-2):

$$G_{k,ex} = \sum_{Z>0} \sum_{IZ} \wp_{IZ}(Y) \cdot L_{IZ} \tag{5.4-1}$$

where *IZ* is the *component array* that defines the excess interactions, and $\wp_{IZ}(Y)$ stands for the corresponding product of site fractions from the *Y* matrix. The interaction parameters L_{IZ} depend on temperature and pressure, and may depend on composition, which is usually represented as Redlich-Kister polynomials of degree $r \leq 3$:

$$L_{I1(Y)} = \sum_{r=0}^{N} {}^{r}L_{I1} \left(y_{s,m1} - y_{s,m2}\right)^{r} \tag{5.4-2}$$

where $m1$ and $m2$ are indices of two moieties on sublattice s according to the component $I1$. The adopted T,P dependence is the same as in the multicomponent Redlich-Kister model:

$$ {}^{r}L_{I1} = a + b \cdot T + c \cdot \ln T + d \cdot P \tag{5.4-3} $$

The compositional dependence eq (5.4-2) is used mainly in single-sublattice models where there is one (or a few) interaction parameter. For instance, in a binary solid solution (A,B)L the $G_{k,ex}$ function eq (5.4-1) with $Z = 1$ takes the form: $G_{k,ex} = y_{1,A}y_{1,B}L_{AB}$, so the Redlich-Kister polynomial expansion to 2nd or 3rd degree is reasonable.

Here and below, in the interaction parameter index, sublattices are separated by colons, e.g. $L_{AB:L}$ means a parameter of interaction between moieties A and B on the first sublattice when the second sublattice is occupied by moiety L, and so on. The order of sublattices is fixed and corresponds to that in the formulae of end members (see section 5.1.1). In each sublattice, only binary interactions are considered. Z can be seen as the parameter order such that $0 < Z \leq \Xi_k$ (the number of sublattices with occurring substitutions).

If mixing occurs in two or more sublattices, the number of parameters will be large enough even if $L_{IZ} = {}^{0}L_{IZ}$ are composition-independent (i.e. regular). For instance, in a reciprocal system (A,B)(L,M), $Z \leq 2$, and the $G_{k,ex}$ function has 5 interaction parameters:

$$G_{k,ex} = y_{1,A}y_{1,B}y_{2,L}L_{AB:L} + y_{1,A}y_{1,B}y_{2,M}L_{AB:M} + y_{1,A}y_{2,L}y_{2,M}L_{A:LM} + y_{1,B}y_{2,L}y_{2,M}L_{B:LM} + y_{1,A}y_{1,B}y_{2,L}y_{2,M}L_{AB:LM} \tag{5.4-4}$$

In (A,B,C)(L,M) system, $Z \leq 2$, but there are 3 moieties substituting in the first sublattice, so $G_{k,ex}$ has 12 parameters:

$$G_{k,ex} = y_{1,A}y_{1,B}y_{2,L}L_{AB:L} + y_{1,A}y_{1,C}y_{2,L}L_{AC:L} + y_{1,B}y_{1,C}y_{2,L}L_{BC:L} + y_{1,A}y_{1,B}y_{2,M}L_{AB:M} + y_{1,A}y_{1,C}y_{2,M}L_{AC:M} + y_{1,B}y_{1,C}y_{2,M}L_{BC:M} + y_{1,A}y_{2,L}y_{2,M}L_{A:LM} + y_{1,B}y_{2,L}y_{2,M}L_{B:LM} + y_{1,C}y_{2,L}y_{2,M}L_{C:LM} + y_{1,A}y_{1,B}y_{2,L}y_{2,M}L_{AB:LM} + y_{1,A}y_{1,C}y_{2,L}y_{2,M}L_{AC:LM} + y_{1,B}y_{1,C}y_{2,L}y_{2,M}L_{BC:LM} \tag{5.4-5}.$$

In an (A,B)(L,M)(X,Y) system, $Z \leq 3$, and $G_{k,ex}$ function has 19 interaction parameters:

$$G_{k,ex} = y_{1,A}y_{1,B}y_{2,L}y_{3,X}L_{AB:L:X} + y_{1,A}y_{1,B}y_{2,M}y_{3,X}L_{AB:M:X} + y_{1,A}y_{2,L}y_{2,M}y_{3,X}L_{A:LM:X} + y_{1,B}y_{2,L}y_{2,M}y_{3,X}L_{B:LM:X} + y_{1,A}y_{1,B}y_{2,L}y_{3,Y}L_{AB:L:Y} + y_{1,A}y_{1,B}y_{2,M}y_{3,Y}L_{AB:M:Y} + y_{1,A}y_{2,L}y_{2,M}y_{3,Y}L_{A:LM:Y} + y_{1,B}y_{2,L}y_{2,M}y_{3,Y}L_{B:LM:Y} + y_{1,A}y_{2,L}y_{3,X}y_{3,Y}L_{A:L:XY} + y_{1,B}y_{2,L}y_{3,X}y_{3,Y}L_{B:L:XY} + y_{1,A}y_{2,M}y_{3,X}y_{3,Y}L_{A:M:XY} + y_{1,B}y_{2,M}y_{3,X}y_{3,Y}L_{B:M:XY} + y_{1,B}y_{2,L}y_{2,M}y_{3,X}y_{3,Y}L_{B:LM:XY} + y_{1,A}y_{2,L}y_{2,M}y_{3,X}y_{3,Y}L_{A:LM:XY} + y_{1,A}y_{1,B}y_{2,M}y_{3,X}y_{3,Y}L_{AB:M:XY} + y_{1,A}y_{1,B}y_{2,L}y_{3,X}y_{3,Y}L_{AB:L:XY} + y_{1,A}y_{1,B}y_{2,L}y_{2,M}y_{3,Y}L_{AB:LM:Y} + y_{1,A}y_{1,B}y_{2,L}y_{2,M}y_{3,X}L_{AB:LM:X} + y_{1,A}y_{1,B}y_{2,L}y_{2,M}y_{3,X}y_{3,Y}L_{AB:LM:XY} \tag{5.4-6}.$$

One additional coefficient per parameter for the temperature dependence would lead to 38 parameter coefficients to fit. This shows that, in multi-sublattice models, L_{IZ} should be treated as (pseudo)regular interaction parameters that are independent of composition. In two-sublattice models, if a lot of good-quality data is available for parameter fitting, subregular $L_{IZ} = {}^{0}L_{IZ} + {}^{1}L_{IZ}(y_{s,m1} - y_{s,m2})$ can be considered as an extraordinary option.

The chemical potential of j-th end member can be obtained as a partial derivative of eq (5.3-3) on the amount of j-th end member (eq. 5.3-6) [1981SUN/ÅGR]. Because of the additivity of G_k function eq (5.3-2), one can write for the partial excess Gibbs energy (activity coefficient) of j-th end member:

$$RT\ln\gamma_j = G_{k,ex} + \sum_s \left(\frac{\partial G_{k,ex}}{\partial y_{s,m_j}} - \sum_m y_{s,m} \frac{\partial G_{k,ex}}{\partial y_{s,m}} \right), \ s \in \Xi_k, m \in M_k \tag{5.4-7}$$

where

$$\frac{\partial G_{k,ex}}{\partial y_{s,m_j}} = \sum_{Z>0} \sum_{JZ} \wp_{JZ}(Y) \cdot \frac{\delta_{J,s,m_j}}{y_{s,m}} \left(L_{JZ} + y_{s,m} \frac{\partial L_{JZ}}{\partial y_{s,m}} \right), \ J \in ?_k \tag{5.4-8}$$

$$\frac{\partial G_{k,ex}}{\partial y_{s,m}} = \sum_{Z>0} \sum_{JZ} \wp_{JZ}(Y) \cdot \frac{\delta_{J,s,m}}{y_{s,m}} \left(L_{JZ} + y_{s,m} \frac{\partial L_{JZ}}{\partial y_{s,m}} \right), \ J \in ?_k \tag{5.4-9}$$

Here, $\delta_{j,s,m}$ is the special form of Kronecker's delta function: $\delta_{j,s,m} = 1$ if $\eta_{j,s,m} \neq 0$, else $\delta_{j,s,m} = 0$. If L_{JZ} is independent of composition then and $\frac{\partial L_{JZ}}{\partial y_{s,m}} = 0$, and two above equations take the form:

$$\frac{\partial G_{k,ex}}{\partial y_{s,m_j}} = \sum_{Z>0} \sum_{JZ} \wp_{JZ}(Y) \cdot \frac{\delta_{J,s,m_j}}{y_{s,m}} L_{JZ}, \ J \in ?_k \tag{5.4-10}$$

$$\frac{\partial G_{k,ex}}{\partial y_{s,m}} = \sum_{Z>0} \sum_{JZ} \wp_{JZ}(Y) \cdot \frac{\delta_{J,s,m}}{y_{s,m}} L_{JZ}, \ J \in ?_k \tag{5.4-11}.$$

If $L_{IZ} = {}^0L_{IZ} + {}^1L_{IZ}(y_{s,m1} - y_{s,m2})$ then …

5.5. Example: (Li,K)(F,Cl) system modelling with CEF

An example of the *CEF* application for the (Li,K)(F,Cl) solid solution system modelling is presented here. Parameters are not optimal and are given below for the sake of demonstration. They are independent of temperature and pressure.

Parameter	J/mol
$L_{\text{Li,K:Cl}}$	5000
$L_{\text{Li,K:F}}$	750
$L_{\text{K:F,Cl}}$	2300
$L_{\text{Li:F,Cl}}$	-820
$L_{\text{Li,K:F,Cl}}$	1200

In this reciprocal system, four end-members are considered:

End member	End-member sublattice formula (as written in DComp record)
LiF	{Li}:{F}:
LiCl	{Li}:{Cl}:
KF	{K}:{F}:
KCl	{K}:{Cl}:

The end member moiety-multiplicity table looks as follows:

m	0	1	2	3	
j	{Li}0*	{F}1	{K}0	{Cl}1	End member
0	1	1	0	0	LiF
1	0	1	1	0	KF
2	1	0	0	1	LiCl
3	0	0	1	1	KCl

*The digit indicates the index of sublattice.

Equation (5.4-1) expands into the form similar to eq (5.4-4):

$$G_{k,ex} = y_{0,Li}y_{0,K}y_{1,F}L_{Li,K:F} + y_{0,Li}y_{0,K}y_{1,Cl}L_{Li,K:Cl} + +y_{0,Li}y_{1,F}y_{1,Cl}L_{Li:F,Cl} + y_{0,K}y_{1,F}y_{1,Cl}L_{K:F,Cl}$$
$$+ y_{0,Li}y_{0,K}y_{1,F}y_{1,Cl}L_{Li,K:F,Cl}$$

To set up five interaction parameters from eq (5.4-1), remake the Phase record and set the CEF ('$') built-in ('S') model with four ipxT columns for indices and the total number of input interaction parameters (in this example, 5). In Page 3 of the Phase window, you will see a combined list filled with zeros. Click "Calculate" toolbar button to see the end member moiety-multiplicity table and the indexes of moieties (columns) in that table.

As shown below, for each interaction parameter, the permutation of moieties d,e (in sublattice 0) and f,g (in sublattice 1) requires four ipxT columns, i.e. two indices per sublattice in order to depict pair-wise interactions. Therefore, interaction parameters given above ($L_{Li,K:Cl}$, $L_{Li,K:F}$, $L_{Li:F,Cl}$, $L_{K:F,Cl}$, $L_{Li,K:F,Cl}$) should be set as shown below:

e (0)*	f (0)*	g (1)*	h (1)*	a**	b	c	d
ipxT[0]	ipxT[1]	ipxT[2]	ipxT[3]	ph_cf[0]	ph_cf[1]	ph_cf[2]	ph_cf[3]
0	2	1	-1	750	0.0	0.0	0.0
0	2	3	-1	5000	0.0	0.0	0.0
0	-1	1	3	2300	0.0	0.0	0.0
2	-1	1	3	-820	0.0	0.0	0.0
0	2	1	3	1200	0.0	0.0	0.0

*The digit indicates the index of sublattice.

**Coefficients a, b, c, d describe T,P dependence (eq 5.4-3) of interaction parameter.

The unused index (all parameters but the last one) is set as -1 in the respective ipxT column.

Note that, unlike in the Berman model, no ipxT column with index of sublattice is used in CEF setup. This is possible because this index can be extracted from the moiety index (i.e. the column index in the moiety-multiplicity table).

5.5. *Modified Bragg-Williams Model*

Modified Bragg-Williams (MBW) model represents a modification of the Compound Energy Formalism (CEF) that allows for more complex description of the Gibbs energy surface to emulates both short-range (SRO) and long-range (LRO) ordering (Liu et al., 2018).

In MBW model, as well as in CEF model, the total Gibbs energy G_k per mole of k-th phase is represented by eq (5.3-3), $G_k = G_{k,ref} - TS_{k,id}^{conf} + H_{k,mix}$. The reference frame term, $G_{k,ref}$, and the configurational entropy, are expressed respectively by equations:

$$G_m^{ref} = \sum_I G_I^o \wp_I(Y) \tag{5.5-1}$$

$$S_{k,id}^{conf} = -R \sum_s \sum_j n^s y_j^s \ln(y_j^s) \tag{5.5-2}$$

Both terms are the same as considered above for Berman and for CEF approaches. The enthalpy of mixing term for the two-sublattice case with two primary end members is expressed by eq (5.5-3):

$$
\begin{aligned}
H_m^{mod} = x_1 x_2 &\left(x_1 \left(L_{A:B,A}^2 + L_{A:B,A}^1 + 2\Delta H_{A:B}^0 \right) + x_2 \left(L_{A:B,B}^1 + L_{A:B,B}^2 + 2\Delta H_{A:B}^0 \right) \right) \\
&+ x_1 x_3 \left(x_1 L_{A:B,A}^2 + x_3 L_{A:B,A}^2 \right) + x_2 x_3 \left(x_2 L_{A:B,B}^1 + x_3 L_{A:B,B}^1 \right) + x_3 \Delta H_{A:B}^0 \\
&+ x_1 x_2 x_3 \left(L_{A:B,A}^2 + L_{A:B,B}^2 + L_{A:B,A}^1 + L_{A:B,B}^1 + 2\Delta H_{A:B}^0 \right)(1 - \sigma)
\end{aligned}
\tag{5.5-3}
$$

where the parameters are described via the first-order Redlich-Kister polynomials

$$L_{1,A:B} = L_{1,0,A:B} + \left(y_{1,A} - y_{1,B} \right) L_{1,1,A:B} \tag{5.5-4a}$$

$$L_{2,A:B} = L_{2,0,A:B} + \left(y_{2,A} - y_{2,B} \right) L_{2,1,A:B} \tag{5.5-4b},$$

$\Delta H_{A,B}^0$ represents the free energies of ordered and anti-ordered compounds, and σ is a correction term, which helps to adjust the level of stabilization of intermediate states by a scaling parameter in the range from 0.0 to 1.0.

Site fractions and ternary end-member mole fractions are related by the equation:

$$y_A^1 = x_1 + x_3; y_A^2 = x_1; y_B^1 = x_2; y_B^2 = x_2 + x_3 \tag{5.5-5}.$$

In GEM-Selektor, interaction parameters for the MBW model can be expressed in two different forms:

$$y_{1,A} y_{1,B} \left(y_{1,A} - y_{1,B} \right) L_{A:B} \quad \text{or} \quad y_{1,A} y_{1,B} L_{A:B},$$

set as type 0 and 1 respectively in the column "type" of the interaction parameter table. This simplifies the description of interaction parameters in GEM-Selektor, while still keeping it quite flexible.

To describe the interaction between moieties in the MBW model, we only have to indicate the indices of the involved moieties and the corresponding L parameter. The respective sublattice indexes are defined by the moiety indexes during the calculations.

As mentioned above, the MBW model allows for two types of interaction parameters: type 0 is given in the form $y_e y_f y_g y_h y_i L$, and type 1 is given in the form $y_e y_f y_g (y_h - y_i) L$. To skip the moiety and its site fraction, the index should just be set to -1 in index table.

Example table:

index	Moiety indexes					type		IP coefficients		
p	e	f	g	h	i	j		a	b	c
0	0	2	-1	-1	-1	0		$L_{A:B}$	0.0	0.0
1	0	2	-1	0	2	1		$L_{A:B}$	0.0	0.0
2	0	2	1	0	2	1		...	...	...

In this case, the first interaction parameter (p=0) between moieties 0 and 2 is expressed in the form $y_0 y_2 L_{0:2}$. The second interaction parameter (p=1) is defined in the form $y_0 y_2 (y_0 - y_2) L_{0:2}$. And the last interaction parameter (p=2) is defined as $y_0 y_2 y_1 (y_0 - y_2) L_{0:2}$.

5.5.1 Example (CaCO₃ – MgCO₃ system)

The MBW model definition for the *CaCO₃ – MgCO₃* system is presented below. The CEF parameters for the 2-sublattice model of calcite-magnesite solid solution were computed with the single defect method, cf. (Liu et al., 2018).

Parameter	kJ/mol
$L_{A:B,B}^{1,0}$	20.595
$L_{A:B,B}^{1,1}$	-0.665
$L_{A:B,A}^{2,0}$	16.21
$L_{A:B,A}^{2,1}$	-0.75
$G_{A,A}^{0}$	0
$G_{B,B}^{0}$	0
$G_{A,B}^{0} = \Delta H_{A,B}^{0}$	-4.12
$G_{B,A}^{0} = \Delta H_{A,B}^{0}$	-4.12

Four end-members need to be considered in this example:

End member	End-member formula (as written in DComp record)
CaCO3 double	{Ca}:{Ca}:[(CO3)2]
CaMg(CO3)2 ordered	{Mg}:{Ca}:[(CO3)2]
CaMg(CO3)2 anti-ordered	{Ca}:{Mg}:[(CO3)2]
MgCO3 double	{Mg}:{Mg}:[(CO3)2]

The corresponding end member- moiety multiplicity table will look like

m	0	1	2	3	GEMS record key		
j	{Ca}0	{Ca}1	{Mg}0	{Mg}1			
0	1	1	0	0	s	CaCO	CaCO3
1	0	1	1	0	s	CaMgCO	CaMgCO3–Ant
2	1	0	0	1	s	CaMgCO	CaMgCO3–Ord
3	0	0	1	1	s	MgCO	MgCO3

Here, the number to the right of the moiety ({Ca}<u>0</u>) is the index of the sublattice for this particular moiety.

The interaction parameters should be expressed in the form accepted by GEMS-Selector package, as it was described above. The whole table (to be entered in Page 3 of the Phase window) is shown below.

Comment	e	f	g	h	i	j	a	b	c	d
dataobj	ipxT[0]	ipxT[1]	ipxT[2]	ipxT[3]	ipxT[4]	ipxT[5]	ph_cf[0]	ph_cf[1]	ph_cf[2]	ph_cf[3]
Lab10	0	2	3	−1	−1	0	41190	0	0	0
Lab11	0	2	3	0	2	1	−1330	0	0	0
Lab20	1	3	0	−1	−1	0	32420	0	0	0
Lab21	1	3	0	1	3	1	−1500	0	0	0
Lab20	0	2	1	−1	−1	0	32420	0	0	0
Lab21	0	2	1	0	2	1	−1500	0	0	0
Lab10	1	3	2	−1	−1	0	41190	0	0	0
Lab11	1	3	2	1	3	1	−1330	0	0	0
DHab0	0	3	−1	−1	−1	0	−8246	0	0	0
DHab0	1	2	−1	−1	−1	0	−8246	0	0	0
KW123	2	1	−1	0	1	1	−130728	0	0	0
KW123	2	1	0	0	1	1	2830	0	0	0
KW123	2	1	2	0	1	1	−2830	0	0	0
KW123	2	1	1	0	1	1	2830	0	0	0
KW123	2	1	3	0	1	1	−2830	0	0	0

References

Abrams D.S., Prausnitz J.M. (1975) Statistical thermodynamics of liquid mixtures: A new expression for the excess Gibbs energy of partly or completely miscible systems. *Amer. Inst. Chem. Eng. J.,* 21, 116-128.

Anderson G.M., Crerar D.A. (1993) *Thermodynamics in geochemistry: The equilibrium model.* Oxford Univ. Press, New York, 588 p.

Anderson G.M. (2006) *Thermodynamics of natural systems.* Second ed., Cambridge University Press, Cambridge, 648 p.

Aranovich L.Y. (1991) *Mineral equilibria of multi-component solid solutions.* Moscow, Nauka, 252 p. (in Russian).

Berman R.G. (1988) Internally-consistent thermodynamic data for minerals in the system Na_2O-K_2O-CaO-MgO-FeO-Fe_2O_3-Al_2O_3-SiO_2-TiO_2-H_2O-CO_2. *Journal of Petrology* **29**(2), 445-522.

Borisov M.V., Shvarov Y.V. (1992) *Thermodynamics of geochemical processes.* Moscow, MSU Publ., 254 p. (in Russian).

Evans K.A., Powell R., Holland T.J.B. (2010) Internally-consistent data for sulphur-bearing phases and application to the construction of pseudosections for mafic greenschist facies rocks in NCKFMASCOSH. *J. Metam. Geol.* (in press)

Glynn P. (2000) Solid solution solubilities and thermodynamics: Sulfates, carbonates and halides. *Rev. Mineral. Geochem.*, **40,** 481-511.

Grenthe I., Puigdomenech I. (1997) *Modelling in aquatic chemistry*, NEA OECD, Paris, 724 p.

Helffrich G., Wood B. (1989) Subregular model for multicomponent solutions. *Amer. Mineral.,* **74**, 1016-1022.

Hillert M. (1998) *Phase equilibria, phase diagrams and phase transformations.* Cambridge University Press, Cambridge, 538 p.

Hillert M. (2001) The compound energy formalism. *J. Alloys and Compounds*, **320**, 161-176.

Holland T.J.B., Powell R. (2003) Activity-composition relationships for phases in petrological calculations: an asymmetric multicomponent formulation. *Contrib. Mineral. Petrol.*, **145**, 492-501.

Holland T.J.B., Powell R. (1996) Thermodynamics of order-disorder in minerals. 2. Symmetric formalism applied to solid solutions. *Amer. Mineral.*, **81**, 1425-1437.

Karpov I.K., Chudnenko K.V., Kulik D.A. (1997) Modeling chemical mass-transfer in geochemical processes: Thermodynamic relations, conditions of equilibria and numerical algorithms. *Amer. J. Sci.,* **297**, 767-806.

Karpov I.K., Chudnenko K.V., Kulik D.A., Avchenko O.V., Bychinski V.A. (2001) Minimization of Gibbs free energy in geochemical systems by convex programming. *Geochem. Internat.*, **39**, 1108-1119.

Liu X., Vinograd V.L., Nichenko S., Kulik D.A., Lu X., Winkler B. (2018): Emulation of short-range ordering within the Compound Energy Formalism: Application to the calcite-magnesite solid solution. *Calphad*, in press.

Lukas H. L., Fries S., Sundman B. (2007) *Computational thermodynamics: The Calphad method.* Cambridge, Cambridge University Press.

Nordstrom D.K., Munoz J.L. (1994) *Geochemical thermodynamics.* Palo Alto, CA, Blackwell Scientific.

Oates W.A. (1969) Ideal solutions. *J. Chem. Educ.,* **46**, 501-504.

Parra T., Vidal O., Agard P. (2002) A thermodynamic model for Fe-Mg dioctahedral K white micas using data from phase-equilibrium experiments and natural politic assemblages. *Contrib. Mineral. Petrol.* **143**, 706-732.

Powell R., Holland T.J.B. (1993) On the formulation of simple mixing models for complex phases. *Amer. Mineral.,* **78**, 1174-1180.

Price J.G. (1985) Ideal site mixing in solid solutions, with an application to two-feldspar geothermometry. *Amer. Mineral.,* **70**, 696-701.

Renon H., Prausnitz J.M. (1968) Local compositions in thermodynamic excess functions for liquid mixtures. *AIChE J.,* **14**, 135-144.

Sundman B., Ågren J. (1981) A regular solution model for phases with several components and sublattices, suitable for computer applications. *J.Phys.Chem.Solids* 42, 297-301.

Wood B.J., Nicholls J. (1978) The thermodynamic properties of reciprocal solid solutions. *Contrib. Mineral. Petrol.* 66, 389-400.

Appendix 5.1. Built-in options for activity coefficients for end members of condensed (solid and liquid) non-electrolyte solution phases (GEM-Selektor v. 3.4).

Code	Model	Equations	Interaction parameters (to enter into `ph_cf` array)
G	Redlich-Kister model (binary)	4.2-36 to 4.2-38	three constant binary interaction parameters
M	Subregular model (binary)	4.2-10 to 4.2-13	two binary interaction parameters, each represented with 3 coefficients as $a - b \cdot T + c \cdot P$
T	Regular model (ternary)	4.2-21 to 4.2-24	one binary interaction parameter per pair and 1 ternary interaction parameter, each represented with 3 coefficients as $a - b \cdot T + c \cdot P$
R	Regular model (multicomponent)	4.2-29, 4.2-30	one binary interaction parameter per pair of components, represented as $a + b \cdot T + c \cdot P$
V	Van Laar model (multicomponent)	4.2-31 to 4.2-35	one binary interaction parameter per pair of components, represented as $a + b \cdot T + c \cdot P$ one size parameter per component (end member) in `dc_cf` array
B	Berman multisite (multicomponent)	5.1-14 to 5.2-6	one binary or ternary interaction parameter for moieties on a sublattice site, represented as $a - b \cdot T + c \cdot P$
$	CEF multisite (multicomponent)	5.1-14; 5.3-6; 5.4-3	one binary, ternary or quaternary interaction parameter for moieties on a sublattice site, represented as $a + b \cdot T + c \cdot T \cdot \ln(T) + d \cdot P$
#	MBW CEF multisite model (both LRO and SRO)	5.5-1 to 5.5-5	one binary, ternary or quaternary interaction parameter for moieties on a sublattice site, represented as $a + b \cdot T + c \cdot T \cdot \ln(T) + d \cdot P$
K	Redlich-Kister (multicomponent)	4.2-41, 4.2-42	four binary interaction parameters per pair of components, represented as $a + b \cdot T + c \cdot T \cdot \ln(T) + d \cdot P$
L	NRTL model (multicomponent)	4.2-50, 4.2-51	two normalized binary interaction energies per pair of components, represented as $A + B/T + C \cdot T + D \cdot \ln(T)$ 1 parameter α per pair of components: $E + F \cdot (T-273.15)$
W	Wilson model (multicomponent)	4.2-59, 4.2-60	two binary interaction energies per pair of components, represented as $\exp[\, A + B/T + C \cdot T + D \cdot \ln(T)\,]$
U	User-defined		Must be provided as Phase scripts and parameters

Appendix 5.2. Modification codes for specific mixing rules (or form of temperature correction) for various EoS and activity (mixing) models in Phase wizard (GEM-Selektor v.3.4).

Code	Displayed message	Comment
N	Default mixing rule or form of interaction parameter coefficients	The appropriate mixing rule or form of interaction parameter coefficients will be selected depending on the chosen fluid- or aqueous model
W	Basic Van der Waals mixing rules in cubic EoS models (default)	Selecting this code enables the automatically assembled gas/fluid phase (chosen in the AutoPhase dialog) because all (default) interaction parameters are zeros
C	Constant one-term binary interaction parameter k_{ij}	One coefficient per non-zero interaction parameter to be entered into ph_cf array
T	Temperature-dependent one-term binary interaction parameter k_{ij}	Two coefficients per non-zero interaction parameter (A_{ij} and B_{ij}) to be entered into ph_cf array
5	5-term Pitzer model parameter temperature dependence	Five coefficients a,b,c,d,e of the temperature dependence $a + b(1/T - 1/298.15) + c \ln(T/298.15) + d(T - 298.15) + e(T - 298.15)^2$. Used also in TOUGHREACT software.
6	6-term Pitzer model parameter temperature dependence	FREZCHEM variant (TBD)
8	8-term Pitzer model parameter temperature dependence	THEREDA variant (TBD)
R	Calculate reciprocal terms (CEF) with 'B','$' and '#' sublattice solution models	All possible end members must be present in the phase definition; no partial site substitution allowed in end members. No additional parameters needed. Can be used as an option with Berman 'B', CEF '$' and MBW '#' sublattice solution models.

Note: Applicable only to non-ideal fluid models 'P', 'E', '7', as well as sublattice solid solution models, also in standalone GEMS3K calculations. Codes 'C' and 'T' require entering non-zero parameter coefficients into ipxT and ph_cf arrays in TSolMod convention, therefore, a Phase record using them must be created in the modelling project before GEM calculations.