

On linear programming approach for the calculation of chemical equilibrium in complex thermodynamic systems

G. Belov

Received: 14 April 2009 / Accepted: 23 July 2009 / Published online: 15 August 2009
© Springer Science+Business Media, LLC 2009

Abstract The advantages of linear programming approach for the calculation of chemical equilibrium in complex thermodynamic systems are demonstrated. The algorithm presented is used for the development of software package for thermodynamic modeling.

Keywords Mathematical modeling · Chemical equilibrium · Thermodynamic systems

1 Introduction

To solve many problems in metallurgy, chemical industry, combustion, gas dynamics, etc. scientists and engineers often use the software for thermodynamic equilibrium calculations [1–8]. Main specific features of the problem of calculation of complex chemical equilibrium from mathematical point of view may be stated as follows [5]:

1. the range of variations of equilibrium concentrations of species may vary from 1 down to very small values of order 10^{-100} ,
2. system of linear equations describing the material balance may be degenerate,
3. chemical potential of species in a mixture is not defined if the phase is absent in equilibrium,
4. one often does not know in advance which phases are present in equilibrium thermodynamic system, so there is the problem of the phase composition determining, and this results in appearance of inequalities in mathematical formulation of the problem.

G. Belov (✉)

Chemistry Department, Moscow State University, 119992 Moscow, Russian Federation
e-mail: gbelov@yandex.ru

May be due to these specific features no reliable algorithm of calculation of complex chemical equilibrium is published up to now.

We try to demonstrate in present study that in many cases the methods of linear programming help to solve the problem mentioned.

2 Problem formulation

Mathematical formulations of the problem of calculation of complex chemical equilibrium and methods of its solution are presented in [1–3]. In most of publications the thermodynamic systems at assigned values of pressure p and temperature T or volume V and T are investigated. Thermodynamic systems of such kind are most often met in solution of practical problems. To simplify the further description (but without loss of generality) let us assume that the volume of condensed phases is negligibly small, and the gas-phase behavior is described by the ideal gas equation of state. When the chemical equilibrium is specified by values of p and T in a thermodynamic system with several possible pure condensed phases and only one mixture (e.g., gas phase) the problem may be stated as follows: find the coordinates of minimum of Gibbs energy

$$G = \sum_{i \in I_C} n_i g_i^\circ + n_{M1} \sum_{i \in I_{M1}} (n_i/n_{M1})(g_i^\circ + \ln(n_i/n_{M1})) \rightarrow \min \quad (1)$$

subject to constraints

$$\sum_{i \in I_T} a_{ji} n_i = b_j, \quad j \in I_E, \quad (2)$$

$$n_{M1} = \sum_{i \in I_{M1}} n_i, \quad (3)$$

$$n_i \geq 0, \quad i \in I_T. \quad (4)$$

Here I_C is a set of indices of condensed species that can form pure phases; I_{M1} is a set of indices of gaseous species; I_E is a set of indices of chemical elements that form the species of thermodynamic system; I_T is a set of indices of all species that form the thermodynamic system; n_i is the equilibrium concentration of i -th species; b_j is the amount of j -th chemical element; a_{ji} is the number of atoms of j -th element in i -th species;

$$g_i^\circ = \mu_i^\circ + \ln(p), \quad i \in I_{M1},$$

$$g_i^\circ = \mu_i^\circ, \quad i \in I_C$$

$$\mu_i^\circ = [H_i^\circ(T) - T S_i^\circ(T)]/(RT),$$

$H_i^\circ(T)$, $S_i^\circ(T)$ are the values of enthalpy and entropy of i -th species at standard pressure (usually 1 bar or atm), that are function of T , and $H_i^\circ(T)$ includes the heat of formation of the species, R is gas constant.

The dual formulation of the problem (1–4) is as follows

$$\sum_{j \in I_E} \lambda_j b_j \rightarrow \max \quad (5)$$

subject to constraints

$$g_i^\circ - \sum_{j \in I_E} a_{ji} \lambda_j \geq 0, \quad i \in I_C, \quad (6)$$

$$\sum_{i \in I_{M1}} \exp \left(-g_i^\circ + \sum_{j \in I_E} a_{ji} \lambda_j \right) \leq 1, \quad (7)$$

$$\lambda_j \leq 0, \quad j \in I_E. \quad (8)$$

Here λ_j are chemical potentials of components, [9].

If the existence of S mixtures in the thermodynamic system examined is possible (I_S is a set of indices of possible mixtures), the Eqs. 1 and 3 look as follows

$$G = \sum_{i \in I_C} n_i g_i^\circ + \sum_{k \in I_S} n_{Mk} \sum_{i \in I_{Mk}} (n_i/n_{Mk})(g_i^\circ + \ln(n_i/n_{Mk})) \rightarrow \min \quad (9)$$

$$n_{Mk} = \sum_{i \in I_{Mk}} n_i, \quad k \in I_S, \quad (10)$$

The condition (7) now should be replaced by the following set of constraints

$$\sum_{i \in I_{Mk}} \exp \left(-g_i^\circ + \sum_{j \in I_E} a_{ji} \lambda_j \right) \leq 1, \quad k \in I_S. \quad (11)$$

The analysis of the problem in both formulations leads to the conclusion that problem solution must satisfy the following conditions

- 1) $G = \sum_{j \in I_E} \lambda_j b_j$;
- 2) for the pure condensed substances with nonzero equilibrium concentrations (i.e. the phase is present in equilibrium), the constraint (6) must be active;
- 3) for the mixtures that are present in equilibrium the constraints (11) must be active.

If the expressions $\sum_{i \in I_{Mk}} \exp(-g_i^\circ + \sum_{j \in I_E} a_{ji} \lambda_j)$, $k \in I_S$ and $\exp(-g_i^\circ + \sum_{j \in I_E} a_{ji} \lambda_j)$, $i \in I_C$ differ from 1, they describe in some sense the probability of emergence of corresponding phases: the nearer to unity is the sum value, the more is the probability of the phase emergence as the conditions of equilibrium (T , p , etc.) vary. This term is called sometimes the thermodynamic activity of the phase and it is designated by letter a . The situation is possible when the activity equals unity but the number of moles of the phase is zero (case of appearing phase).

The problem formulation changes slightly when the equilibrium is described by assigned values of T and V . In this case it is necessary to calculate coordinates of

conditional minimum of Helmholtz energy A , but only the terms related to gas phase should be modified in (1) and (9)

$$A = \sum_{i \in I_C} n_i g_i^\circ + \sum_{k \in I_S} n_{Mk} \sum_{i \in I_{Mk}} (n_i/n_{Mk})(g_i^\circ + \ln(n_i/n_{Mk})) + n_{M1} \ln(n_{M1}) \rightarrow \min, \quad (12)$$

here $g_i^\circ = \mu_i^\circ + \ln(RT/V)$, $i \in I_{M1}$. Constraints remain the same.

3 Algorithm

More than 50 years ago White et al. [10, 11] presented the algorithm of solution of the problem (1–4) that used the methods of linear programming. The idea of their approach was to approximate the nonlinear terms $\alpha_i \ln \alpha_i$ by a piecewise function. One can do that by setting $\alpha_i = n_i/n_{M1}$ and $\beta_i = \alpha_i \ln \alpha_i$. Such an approach permits to determine the phase and chemical composition of thermodynamic system by means of the iterative solution of the following problem

$$\sum_{i \in I_T} g_i^\circ n_i + \sum_{i \in I_{M1}} \sum_{j=1}^L \beta_{ij} y_{ij} \rightarrow \min \quad (13)$$

subject to constraints (2–4) and

$$n_{M1} = \sum_{j=1}^L y_{ij}, \quad i \in I_{M1}, \quad (14)$$

$$n_i = \sum_{j=1}^L \alpha_{ij} y_{ij}, \quad i \in I_{M1}. \quad (15)$$

$$y_{ij} \geq 0, \quad i \in I_{M1}, j = 1, 2, \dots, L. \quad (16)$$

Here L is the number of points in which pieces of straight lines touch the approximated curve, y_{ij} are the additional variables. For the practical purposes it is enough to use only two straight lines so that $L = 3$.

The matrix of coefficients of constraints (2), (3), (14), (15) is presented in Table 1. N is the number of species in thermodynamic system, N_C is the number of pure condensed phases, and m is the number of chemical elements in thermodynamic system. Number of unknowns equals to $4N - 3N_C + 1$, and the number of constraints is $m + 2N + 1$.

As one can see, in such formulation we have the linear programming problem, and it is possible to use some well-known procedure to solve it.

The procedure for the solution of the problem (1–4) presented in [10, 11] looks as follows.

Table 1 Matrix of coefficients of constraints for target function (1)

$a_{11}n_1 + \dots + a_{1,NC}n_{NC}$	$+ \dots$	$+ a_{1N}n_N$	b_1
$a_{21}n_1 + \dots + a_{2,NC}n_{NC}$	$+ \dots$	$+ a_{2N}n_N$	b_2
\dots	\dots		\dots
$a_{m,1}n_1 + \dots + a_{m,NC}n_{NC}$	$+ \dots$	$+ a_{mN}n_N$	b_m
$n_{NC+1} + n_{NC+2} + \dots + n_N$		$-n_{M1}$	0
		$-n_{M1} \quad x_{11} + x_{12} + x_{13}$	0
		$-n_{M1} \quad x_{21} + x_{22} + x_{23}$	0
		\dots	\dots
		$-n_{M1} \quad x_{N,1} + x_{N,2} + x_{N,3}$	0
$-n_{NC+1}$		$\alpha_{11}x_{11} + \alpha_{12}x_{12} + \alpha_{13}x_{13}$	0
$-n_{NC+2}$		$\alpha_{21}x_{21} + \alpha_{22}x_{22} + \alpha_{23}x_{23}$	0
\dots		\dots	\dots
$-n_N$		$\alpha_{N,1}x_{N,1} + \alpha_{N,2}x_{N,2} + \alpha_{N,3}x_{N,3}$	0

1. For each mixture species compute values β_{ij} at three values of α : 0, 0.5 and 1.
2. Find the solution of the problem (13–16), (2–4).
3. For each $i \in I_{M1}$ compute three new values of α so that $\alpha_{\min} \leq n_i \leq \alpha_{\max}$. E.g. if n_i belongs to the interval $[0,0.5]$, then $\alpha_{ij} = 0, 0.25, 0.5$, and if n_i belongs to the interval $[0.5,1]$ then $\alpha_{ij} = 0.5, 0.75, 1$.
4. Compute values of β_{ij} and solve the problem again.

This cycle iterates many times and at each step for each species belonging to the mixture the interval $[\alpha_{\min}, \alpha_{\max}]$ is reduced. The process ends when the interval length becomes less than the prescribed value ε . The more general description of this optimization technique is presented in [12].

The approach presented may be generalized for the case of arbitrary number of mixtures in thermodynamic system. Let the number of possible mixtures in thermodynamic system is S , then the relations (13–16) may be stated as follows

$$\sum_{i \in I_T} g_i^\circ n_i + \sum_{k \in I_S} \sum_{i \in I_{Mk}} \sum_{j=1}^L \beta_{ij} y_{ij} \rightarrow \min \tag{17}$$

Subject to constraints (2), (4) and

$$n_{Mk} = \sum_{i \in I_{Mk}} n_i, \quad k \in I_S, \tag{18}$$

$$n_{Mk} = \sum_{j=1}^L y_{ij}, \quad i \in I_{Mk}, k \in I_S, \tag{19}$$

$$n_i = \sum_{j=1}^L \alpha_{ij} y_{ij}, \quad i \in I_{Mk}, k \in I_S, \tag{20}$$

$$y_{ij} \geq 0, \quad i \in I_{Mk}, j = 1, 2, 3, k \in I_S. \tag{21}$$

The matrix of coefficients of constraints (2), (18–20) is presented in Table 2. Number of unknowns for this case equals to $4N - 3N_C + S$, and the number of constraints equals to $m + S + 2(N - N_C)$. N_1, N_2, \dots, N_S designate number of species in mixtures 1, 2, ..., S .

It should be stressed that such an approach allows to model equilibrium states of partially non-equilibrium thermodynamic systems in most natural way. Concentrations of one or several species or their ratio are supposed to be fixed in such systems. Constraints of such kind may be caused by influence of kinetic factors and are often met in practice of thermodynamic modeling, e.g., see [8]. In general case these linear constraints may be presented in a way similar to (2)

$$\sum_{i \in I_T} a_{ji} n_i = b_j, \quad j \in I_K, \tag{22}$$

Table 2 Matrix of coefficients of constraints for target function (9)

$a_{11}n_1 + \dots + a_{1,NC}n_{NC}$	$+ \dots + a_{1,N}n_N$						b_1
$a_{21}n_1 + \dots + a_{2,NC}n_{NC}$	$+ \dots + a_{2,N}n_N$						b_2
...							...
$a_{m,1}n_1 + \dots + a_{m,NC}n_{NC}$	$+ \dots + a_{m,N}n_N$						b_m
	$\sum_{i \in I_{M1}} n_i$	$-n_{M1}$					0
	$\sum_{i \in I_{M2}} n_i$		$-n_{M2}$				0
	...						0
	$\sum_{i \in I_{MS}} n_i$			$-n_{MS}$			0
		$-n_{M1}$			$\sum_{j=1}^3 x_{1j}$		0
	
		$-n_{M1}$				$\sum_{j=1}^3 x_{N1j}$	0
				...			0
		$-n_{MS}$					0
			
						$\sum_{j=1}^3 x_{NSj}$	0
							...
	$-n_{NC+1}$						0
	...				$\sum_{j=1}^3 \alpha_{1j} x_{1j}$		0
					...		0
	$-n_{NC+N1}$					$\sum_{j=1}^3 \alpha_{N1j} x_{N1j}$	0

							0
	$-n_N$					$\sum_{j=1}^3 \alpha_{NS} x_{NSj}$	0

where I_K is a set of indices of constraints, a_{ji} is a coefficient of j -th constraint for i -th species, b_j is real number.

The most specific feature of the algorithm is the fact that the phase composition is automatically determined in the process of calculation of equilibrium composition of thermodynamic system. One should also mention the simplicity of the algorithm implementation.

It was found in the process of the algorithm fine-tuning that the reliability of determination of the phase composition depends sometimes on the speed of interval reduction. The recommended in [10, 11] double interval reduction does not guarantee the true phase composition of thermodynamic system. So after the reduction of the interval down to some value it is recommended to check the conditions (6) and (11). And if for some phases these conditions are not true the composition should be determined again with smaller interval reduction coefficient (one can use coefficient 1.5 instead of 2).

The main drawback of the algorithm is its laboriousness that quickly grows with the growth of number of substances in mixtures. Apparently it was this drawback that prevented the wide use of this algorithm in practice of thermodynamic calculations up to now.

But one can enhance the performance of the algorithm combining it with some effective procedure of nonlinear optimization. We have used to that aim the approach based on dual formulation of the problem (5–8). But the final system of Equations 22–25 contains only active constraints because it is supposed that the phase composition is already found on the previous step.

$$\sum_{i \in I_C^*} a_{ji} n_i + \sum_{k \in I_S^*} n_{Mk} \sum_{i \in I_{Mk}} a_{ji} \exp \left(-g_i^\circ + \sum_{l \in I_E} a_{li} \lambda_l \right) = b_j, \quad j \in I_E, \quad (23)$$

$$g_i^\circ - \sum_{j \in I_E} a_{ji} \lambda_j = 0, \quad i \in I_C^*, \quad (24)$$

$$\sum_{i \in I_{Mk}} \exp \left(-g_i^\circ + \sum_{j \in I_E} a_{ji} \lambda_j \right) = 1, \quad k \in I_S^*. \quad (25)$$

Here I_C^* is a set of indices of condensed species that form pure phases, I_S^* is a set of indices of mixtures that are present in thermodynamic system or are appearing in it.

4 Results and discussion

To implement the algorithm described above we have used the FORTRAN procedure DSPLP [13], borrowed from the site gams.nist.gov. This procedure was designed for the solution of the problems of large dimension and allows storing matrix coefficients in a compact form. In order to be able to thoroughly test the software we have supplied it with an extensive database on thermodynamic properties of individual substances [14].

Let us look at several examples illustrating the abilities of the algorithm.

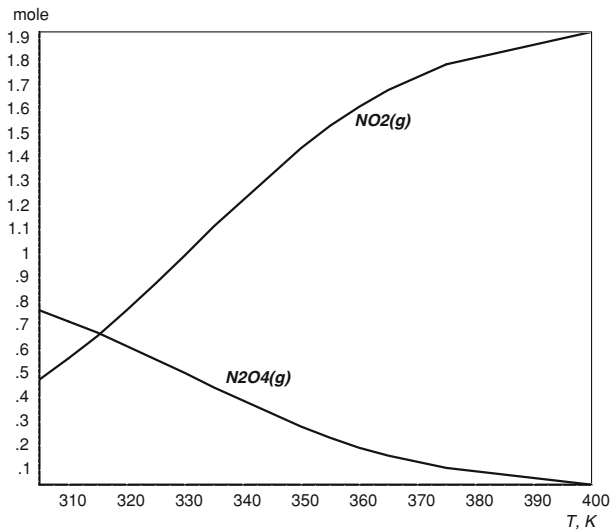


Fig. 1 Equilibrium composition in system $N_2O_4 + 2NO_2$ as function of temperature

1. The algorithm presented works well when the number of reactants in thermodynamic system is equal to (or less than) the number of chemical elements. The material balance equation 2 for the reaction equilibrium $N_2O_4 = 2NO_2$ looks like this

$$\begin{aligned} 2n_{N_2O_4} + n_{NO_2} &= 1, \\ 4n_{N_2O_4} + 2n_{NO_2} &= 2. \end{aligned}$$

As one can see this system of linear equations is degenerated so some special procedure is necessary to compute the chemical equilibrium. But the use of linear programming procedure permits to find the equilibrium composition with the reasonable accuracy, Fig. 1 (assigned pressure is 1 bar). It should be mentioned that this is an example of partial equilibrium. If one tries to model the full chemical equilibrium for this system then the dominating species will be O_2 and N_2 .

2. The algorithm presented provides possibility to reliably identify the situations when assigned pressure can not be achieved due to thermodynamic reasons, because it exceeds the vapor pressure over the condensed phase. In particular, for the thermodynamic system formed by H_2O , calculation at pressure 1 bar and temperature 372 K shows that only liquid water can exist in such conditions at equilibrium (i.e. the vapor pressure is less than 1 bar). To compute partial pressures of the gaseous species one should compute equilibrium composition at assigned values of temperature and volume, or add some inert gas in thermodynamic system, or compute composition at lower pressure. Table 3 presents values of liquid water and the gas phase activities as function of temperature at 1 bar.

3. Sometimes it is necessary to investigate thermodynamic system that is not subject to chemical transformations. But this fact often is not known in advance. For example, the thermodynamic system formed by CO_2 is practically nonreacting up to temperature

Table 3 Activities of water and gas phase as function of temperature at 1bar

T (K)	300	350	373	374	400
$a_{\text{H}_2\text{O}(c)}$	1	1	1	0.9725	0.4188
a_{gas}	0.2853	0.8262	0.9927	1	1

1000K. From the computational point of view that means that both chemical elements are concentrated in one species and the system of materials balance equations 2 is near to degeneracy. Using the linear programming approach one can solve this problem in most natural way. The developed algorithm provides possibility to solve many other problems of such kind. For example using this algorithm one can compute the temperature dependence of thermodynamic properties of some nonreacting mixture such as $\text{CH}_4 + \text{N}_2$.

4. The algorithm of calculation of complex chemical equilibrium described above is very efficient and reliable for computer modeling of various thermodynamic systems including the models with several mixtures.

It appeared to be reasonable to combine the linear programming procedure and the constrained nonlinear optimization algorithms. The time of calculation strongly depends on the number of species in the mixtures of thermodynamic system. The influence of the number of pure phases is significantly lower. This fact is easily explained by the structure of the matrix of constraints—in fact each species in the mixture adds four unknowns and two constraints. One should also mention that the determining of equilibrium composition using the linear programming procedure DSPLP is not always successful, but in most cases it allows determining of the phase composition and it provides a good starting point for further refinement by the nonlinear optimization procedure.

Finally, we should mention that for a long time the linear programming procedures are used in practice of thermodynamic modeling for the calculation of initial approximation [3]. However in this case the calculation is accomplished in an assumption that all species form pure phases so the problem is reduced to calculation of coordinates of constrained minimum of the function

$$G = \sum_{i \in I_T} n_i g_i^\circ \rightarrow \min$$

subject to constraints (2) and (4). The more complicated variant of such an approach was used in [15] for the calculation of the phase diagrams.

Acknowledgments The study was supported by RFBR, grants No. 1 08-03-00221-a and 08-03-00506-a.

References

1. F. Van Zeggeren, S.H. Storey, *The Computation of Chemical Equilibria* (Cambridge University, Oxford, 1970)
2. R. Holub, P. Voňka, *The Chemical Equilibria of Gaseous Systems* (Reidel, Dordrecht, 1976)

3. W.R. Smith, R.W. Missen, *Chemical Reaction Equilibrium Analysis: Theory and Algorithms* (John Wiley, New York, 1982)
4. K. Hack (ed.), *The SGTE Casebook: Thermodynamics at Work* (Taylor & Francis, London, 2007)
5. W.R. Smith, R.W. Missen, Strategies for solving the chemical equilibrium problem and an efficient microcomputer-based algorithm. *Can. J. Chem. Eng.* **66**, 591 (1988)
6. N.R. Amundson et al., Primal-dual active-set algorithm for chemical equilibrium problems related to the modeling of atmospheric inorganic aerosols. *J. Optim. Theory Appl.* **128**, 469 (2006)
7. A. Burrows, C.M. Sharp, Chemical equilibrium abundances in brown dwarf and extrasolar giant planet atmospheres. *Astrophys. J.* **512**, 843 (1999)
8. S.B. Pope, Gibbs function continuation for the stable computation of chemical equilibrium. *Combust. Flame.* **139**, 222 (2004)
9. W.S. Dorn, Variational principles for chemical equilibrium. *J. Chem. Phys.* **32**, 1490 (1960)
10. W.B. White, S.M. Johnson, G.B. Dantzig, Chemical equilibrium in complex mixtures. *J. Chem. Phys.* **28**, 751 (1958)
11. W.B. White, S.M. Johnson, G.B. Dantzig, A linear programming approach to chemical equilibrium problem. *Manage. Sci.* **5**, 38 (1958)
12. G.V. Reklaitis, A. Ravindran, K.M. Ragsdell, *Engineering Optimization: Methods and Applications* (John Wiley, New York, 2006)
13. R.J. Hanson, K.L. Hiebert, A sparse linear programming subprogram, Report SAND81-0297. (Sandia National Laboratories, 1981)
14. G.V. Belov, V.S. Iorish, V.S. Yungman, IVTANTHERMO for Windows—database on thermodynamic properties and related software. *CALPHAD.* **23**, 173 (1999)
15. J.A.D. Connolly, Computation of phase equilibria by linear programming: a tool for geodynamic modeling and its application to subduction zone decarbonation. *Earth Planet. Sci. Lett.* **236**, 524 (2005)