Calphad 61 (2018) 50–61

Contents lists available at ScienceDirect

Calphad

journal homepage: www.elsevier.com/locate/calphad

CpFit program for approximation of heat capacities and enthalpies by Einstein-Planck functions sum

Alexey L. Voskov*, Ilya B. Kutsenok, Gennady F. Voronin†

Department of Chemistry, Lomonosov Moscow State University, 119991 Moscow, Russia

ARTICLE INFO

Keywords:
Heat capacity
Enthalpy
Einstein functions
Model

ABSTRACT

CpFit program for joint approximation of heat capacities and enthalpies (heat contents) by means of Einstein-Planck functions sum has been developed. It uses algorithms for automatic search of number of functions in the sum and automatic search of initial approximation for model parameters. The program also allows to use arbitrary extra terms for approximation of heat capacity anomalies such as lambda-transitions, Schottky anomalies etc. Experimental data for thorium and uranium dioxides, natrolite, potassium and thallium substituted natrolites were successfully approximated in wide ranges of temperature.

1. Introduction

The most common way of approximation of thermodynamic functions of individual substances is using temperature-dependent polynomials [1]. They can be applied for any substances in fairly wide temperature ranges. However, their main drawbacks are problems with extrapolation outside fitted temperature interval and with joint description of low- and high-temperature experimental data. The most common theoretical models of heat capacity, Debye and Einstein-Planck functions, use some simplified assumption about oscillators and their spectrum and usually have low accuracy. And for practical use their combinations with polynomials [2,3] or segmented models [4] were proposed. They have higher accuracy, can be extrapolated down to 0 K and are applied in so-called third generation Calphad databases. However they were tested mainly for pure elements [2–9] and their usage for compounds is less systematic [2,10–18].

The most comprehensive database for $T \leq 298.15$ K that uses the sum of Einstein-Planck function with polynomials was developed by Vřešťál et al. [5] and includes 53 elements. Elements of groups 16, 17, 18, hydrogen, nitrogen, phosphorous, elements with atomic numbers greater than 83 and lanthanides from Ce to Lu were not considered in that work. Models with polynomial terms for Ag, Cu, Mo, Ti, Sn [2], Fe [3,6], Mn [6,7], Co [8], Pb [9] and the segmented model for Cr, Al, Fe [4] were optimized for the entire temperature ranges.

Usage of Debye or Einstein-Planck function together with polynomials for heat capacity and heat content approximation of compounds was reported for $CaCl_2$ [2], $U_3Al_2Si_3$ [10], chromium carbides Cr_3C_2, Cr_7C_3, $Cr_{23}C_6$ [11], NiTi, $Al_{.45}Ni_{.5}Ti_{.05}$, $Al_{.16}Ni_{.74}Ti_{.10}$, $Al_{.25}Ni_{.5}Ti_{.25}$ [12], bismuth silicates $Bi_{12}SiO_{20}$ and $Bi_4Si_3O_{12}$ [13], Al_3Ni [14], HfV_2 [15], plagioclase solid solutions [16]. It should be noted that only for chromium carbides both high- and low-temperature data for 0–2000 K were described by a single combination of Einstein-Planck function with polynomials for each compound. In other cases this model is used either for extrapolation of data for $T > 300$ K to lower temperatures for $S_{298.15}°$ estimation [12,13] or only for description of lower temperature data, i.e. for $T < 300$ K [14,15]. For plagioclase solid solutions Einstein functions were used for 30–298.15 K [16]. For $CaCl_2$ and $U_3Al_2Si_3$ temperature ranges are not given explicitly. There are also examples of using the sum of Debye function, Einstein function and polynomials for some compounds, e.g. calcium titanate $CaTiO_3$ [17] and zirconolite $CaZrTi_2O_7$ [18] for $T < 400$ K.

An alternative approach proposed by Voronin et al. [19] is using Einstein-Planck functions sum for approximation of heat capacity and heat content data in the entire temperature range using empirical parameters set found by the least squares method. It was successfully applied for description of heat capacities of many substances [9,20–23] including ones which undergo second order phase transitions [22]. An approximation of experimental data for solid Pb made by Khvan et al. [9] is especially interesting because both the model from [19] and the sum of Einstein-Planck function with polynomials [2] were used for it. And the calculated values of $C_{p,298.15}°$ and $S_{298.15}°$ for these models are in agreement with each other. However no examples of applying Einstein-Planck functions sum to heat content experimental data approximation are reported in the literature. Such simultaneous usage of both kinds of data may be essential if they cover different temperature intervals, e.g. C_p are obtained for low temperatures and $H(T_2) - H(T_1)$ — for high

* Corresponding author.
E-mail addresses: alvoskov@gmail.com, voskov@td.chem.msu.ru (A.L. Voskov).
† Deceased 15 July 2017.

https://doi.org/10.1016/j.calphad.2018.02.001
Received 1 November 2017; Received in revised form 30 January 2018; Accepted 1 February 2018

temperatures. Approximations of experimental data in [9] were made by a preliminary version of CpFit computer program developed in this work and in [20–23] — by CpAppro computer program. Both software packages were developed in the Laboratory of Chemical Thermodynamics of Lomonosov Moscow State University (see http://td.chem.msu.ru for details). In the case of heat capacity anomalies additional custom MATLAB scripts were used [22]. Such scripting may be a complex task for non-programmers and simple approaches and software are required to effectively apply an approach described above. It should be noted that CpAppro doesn't approximate heat content, heat capacity anomalies and doesn't allow to customize statistical weights of experimental points. Such situations make a development of a new software and algorithms for practical application of Einstein-Planck functions sum model an actual task.

It is also possible to combine Einstein-Planck functions sum with Debye function and polynomials. Such approach was applied by Voronin et al. [19] and by Jankovský et al. [24,25] for heat capacities of complex oxides. However in [19] it was shown that addition of Debye function to Einstein-Planck functions sum usually doesn't improve approximation accuracy. The reason of this result is a possibility to approximate continuous vibrational spectrum of Debye model by a discrete frequencies of Einstein-Planck functions sum [19]. It is an important result because Debye function cannot be expressed through elementary functions and has much higher computational cost and much more complex software implementation than Einstein-Planck function.

In this work CpFit computer program for heat capacity and heat content experimental data approximation was developed. It uses Einstein-Planck functions sum together with arbitrary extra terms to describe $C_p(T)$ anomalies, e.g. second order phase transitions. An universal algorithm for automatic selecting the number of functions in the sum and automatic searching of initial approximation of parameters was also proposed and implemented.

2. Approximation method

An approach suggested in this work and implemented in CpFit program consists of three key steps:

1. Separation of experimental data into baseline data set and anomalies data sets. It is done manually by setting minimal and maximal temperatures $T_{\min}^{(i)}$ and $T_{\max}^{(i)}$ for each anomaly data set i. Baseline data set includes only data that don't involve any phase transitions or other heat capacity anomalies. Each anomaly data set represents a region of phase transition or other C_p anomaly that can't be described solely by Einstein-Planck functions. Note: anomalies data sets can be absent but a baseline data set is always required and can include all points.
2. Approximation of baseline data by Einstein-Planck functions sum by means of the least squares method. The result is $\vec{\alpha}$ and $\vec{\theta}$ parameters vectors (see below for detailed description).
3. Approximation of anomalies data sets by sums of the baseline model and extra term models, i.e.:

$$C_p\left(T, \vec{\alpha}, \vec{\theta}, \vec{\xi}^{(1)}, ..., \vec{\xi}^{(r)}\right) = C_p^{\mathrm{BL}}\left(T, \vec{\alpha}, \vec{\theta}\right) + \sum_{i=1}^{r} C_p^{\mathrm{ex},(i)}\left(T, \vec{\xi}^{(i)}\right), \tag{1}$$

where r is the number of extra terms, $\vec{\xi}^{(i)}$ and $C_p^{\mathrm{ex},(i)}$ are the model parameters vector and extra term function for i-th anomaly data set respectively.

In this case $\vec{\alpha}$ and $\vec{\theta}$ values are fixed (taken from the Step 2) and $\vec{\xi}^{(i)}$ are found by the least squares method (see below for detailed description).

As it was mentioned above, $T_{\min}^{(i)}$ and $T_{\max}^{(i)}$ values are borders of temperature region of i-th heat capacity anomaly. Although an exact technique of their manual search depends on an anomaly shape and experimental data accuracy, the next procedure can be recommended for lambda-transitions:

1. Visually detect anomalies borders and set corresponding $T_{\min}^{(i)}$ and $T_{\max}^{(i)}$ values manually.
2. Evaluate $\vec{\alpha}$ and $\vec{\theta}$ parameters using baseline data.
3. Try to extend borders of anomalies, i.e. decrease $T_{\min}^{(i)}$ and/or increase $T_{\max}^{(i)}$, and repeat step 2 while baseline data approximation error decreases.

In CpFit program optimal parameters values are found by Levenberg-Marquardt method implemented in `levmar` library [26]. This is a numeric iterative method for solving the non-linear least squares problem. Objective functions minimized both in baseline and excess terms case can be represented as:

$$\chi^2(\vec{\beta}) = \sum_{i=1}^{n} W_i^2(\Delta F_i(\vec{\beta}))^2 = \sum_{i=1}^{n} W_i^2(F_i^{\mathrm{calc}}(\vec{\beta}) - F_i^{\mathrm{exp}})^2, \tag{2}$$

where n is the number of experimental points, $\vec{\beta}$ is the vector of the optimized model parameters, W_i and ΔF_i are statistical weight and minimized absolute deviation for i-th point, F_i^{calc} and F_i^{exp} are calculated and experimental function values respectively.

Standard deviations of the β_i parameters are calculated using the next formula:

$$\vec{s}_\beta^{\,2} = \hat{\sigma}^2 \mathrm{diag}[(J^{\top}J)^{-1}];\ J_{ij} = \frac{\partial F_i^{\mathrm{calc}}}{\partial \beta_j};\ \hat{\sigma}^2 = \frac{\chi^2}{n-k}, \tag{3}$$

where J is Jacobian, n is the number of experimental points, k is the number of parameters in the $\vec{\beta}$ vector, $\hat{\sigma}$ is the model standard error, χ^2 is the sum of squares calculated using Eq. (2).

Standard deviations of functions F_i that can be used to estimate their confidence intervals using the following equation:

$$s_Y(T) = \sqrt{J_x^{\top} C J_x} = \hat{\sigma}\sqrt{J_x^{\top}(J^{\top}J)^{-1}J_x}, \tag{4}$$

where C is a covariance matrix, $J_x = J_x(T)$ is a column vector that has the next layout:

$$J_x(T) = \left(\frac{\partial F(\vec{\beta}, T)}{\partial \beta_1} \quad \cdots \quad \frac{\partial F(\vec{\beta}, T)}{\partial \beta_k}\right)^{\top}. \tag{5}$$

Standard deviations of parameters and functions values can be converted to confidence intervals by means of the following equation:

$$\Delta\beta = s_\beta \cdot t_{f,\alpha};\ f = n - k, \tag{6}$$

where $t_{f,\alpha}$ is t-distribution two-sided quantile with probability p (CpFit uses $\alpha = 0.95$ for 95% confidence intervals), $f = n - k$ is the number of freedom degrees. Usually $t_{f,\alpha} \approx 2 \div 3$. Eqs. (4), (5) and (6) were taken from [27].

Implementation of this model in CpFit hides all analytical expressions and technical details and allows to use it as an ordinary statistical package for a non-linear regression. The method of automatic selection of the number of terms in Einstein-Planck functions sum allows to obtain accurate models in most cases.

The developed program has user-friendly graphical interface that allows to visualize data in two-dimensional plots and calculate tabulated values of thermodynamic functions. All results, i.e. parameters sets, plots and functions values, can be exported to graphical files, MS Excel and MATLAB. CpFit program and its source code is distributed under GNU GPL 2 license as free software and can be obtained on the web-site of the Laboratory of Chemical Thermodynamics of Lomonosov Moscow State University (http://td.chem.msu.ru).