Journal of Thermal Analysis and Calorimetry, Vol. 87 (2007) 2, 595–599

CONSTRUCTING A SELF-CONSISTENT INTEGRAL BASELINE BY USING CUBIC SPLINES

D. V. Malakhov[*] *and M. K. Abou Khatwa*

Department of Materials Science and Engineering, McMaster University, Hamilton, Canada

Various types of transformations require different baselines reflecting specificities of these transitions. The present work deals with the case when a degree of transformation is directly proportional to heat consumed or released. For such case, a baseline is named an integral baseline and is traditionally constructed by unnecessary simplifications. A new method is proposed as an alternative fast and robust computational method for baseline construction utilizing interpolating cubic splines. The method is self-consistent in the sense that it is free of needless assumptions and that it provides linearity between the degree of transformation and heat measured.

Keywords: *baseline, cubic splines, DSC*

The essence of a new method

Let us assume, for the sake of determinacy, that an endothermic transformation is studied by using differential power scanning calorimetry and that a time dependence of heat flux is recorded, i.e. that the experiment results in the table t_i, $\dot{Q}_i$, i=1,...,n, where n is a total number of measurements. Usually, an objective of the calorimetric experiment outlined above is extracting the enthalpy of transformation from raw data. Before diving into mathematical details of the new method, let us recall the nature of the signal measured. There are two thermally equivalent containers inside a calorimeter. One of them is filled with a reference material, which does not undergo a phase transformation within an investigated temperature region. The other container is filled with a sample, in which a transformation occurs. Both containers are heated with a constant heating rate α in such a manner that their temperatures remain the same at all times. A time-depending difference $\dot{Q}_S - \dot{Q}_R$ between powers supplied to the sample and to the reference is measured and recorded. Since the reference does not experience the phase transition, energy provided is exclusively spent for heating. In the case of the sample, a portion of energy supplied is still spent for heating, while the rest of it is used for driving the transformation:

$$\dot{Q}_R \, \mathrm{d}t = C_{pR} \, m_R \, \mathrm{d}T \tag{1}$$

$$\dot{Q}_S \mathrm{d}t = C_{pS} m_S \mathrm{d}T + \mathrm{d}H \tag{2}$$

Since the heating rate is constant, temperature increases with time as $T=T_0+\alpha t$, where T_0 is temperature at the beginning of the experiment. Substituting $\mathrm{d}T/\mathrm{d}t=\alpha$ in expressions (1) and (2) yields:

$$\dot{Q}_R = C_{pR} \, m_R \, \alpha \equiv g_R \tag{3}$$

$$\dot{Q}_S = C_{pS} m_S \alpha + \mathrm{d}H / \mathrm{d}t \equiv g_S + \mathrm{d}H / \mathrm{d}t \tag{4}$$

Subtracting (3) from (4) and rearranging gives:

$$\mathrm{d}H / \mathrm{d}t = (\dot{Q}_S - \dot{Q}_R) - (g_S - g_R) \equiv \dot{Q} - g$$

If the transformation starts at time t_{beg} and is completed at time t_{end}, then the enthalpy of transformation can be calculated as:

$$H = \int_{t_{beg}}^{t_{end}} (\dot{Q} - g) \mathrm{d}\tau \tag{5}$$

Expression (5) is not easy to apply. Firstly, integration limits are not known exactly. Secondly, while heat capacities of many substances (e.g. Fe) are known, it cannot be taken for granted that they are known for many other compounds and alloys (e.g. steels). Even if temperature dependencies of heat capacities of the sample's states prior to the transformation and after it are known and even though a method of extrapolating them is reliable, nothing is known about fractions of these states during the transformation. Subsequently, the baseline $g \equiv g_S - g_R$ seen in the right-hand-side of (5) is not readily defined. So how can this elusive baseline be constructed? This question is not of academic interest only. An importance of constructing a 'good baseline' has been accentuated in literature. Examples are known [1] where the magnitude of the enthalpy of reaction was altered by approximately 10% by choosing an improper baseline.

For answering this question, it should firstly be recalled that $\mathrm{d}H/\mathrm{d}t=0$ if $t \le t_{beg}$, which means that the position of the baseline prior to the transformation coincides

* Author for correspondence: malakhov@mcmaster.ca

1388–6150/$20.00

Akadémiai Kiadó, Budapest, Hungary
Springer, Dordrecht, The Netherlands

with $\dot{Q}$. If $t \geq t_{end}$, then $dH/dt=0$ and the location of the baseline after the transition is indistinguishable from $\dot{Q}$.

Secondly, since an extent of the transformation at any time between t_{beg} and t_{end} is deemed directly proportional to energy consumed for driving the transformation by this time, then contributions from the initial and final states to the baseline can be related to a degree of the transformation [2]:

$$g(t)=\left\{1-\frac{\int_{t_{beg}}^{t}[\dot{Q}(\tau)-g(\tau)]d\tau}{\int_{t_{beg}}^{t_{end}}[\dot{Q}(\tau)-g(\tau)]d\tau}\right\}g(t_{beg})+$$

$$+\frac{\int_{t_{beg}}^{t}[\dot{Q}(\tau)-g(\tau)]d\tau}{\int_{t_{beg}}^{t_{end}}[\dot{Q}(\tau)-g(\tau)]d\tau}g(t_{end})= \quad (6)$$

$$=g(t_{beg})+\frac{g(t_{end})-g(t_{beg})}{\int_{t_{beg}}^{t_{end}}[\dot{Q}(\tau)-g(\tau)]d\tau}\int_{t_{beg}}^{t}[\dot{Q}(\tau)-g(\tau)]d\tau$$

However, Eq. (6) cannot be solved with respect to $g(t)$ merely because it is not even clear what class of function the solution should be sought within.

A discrete nature of calorimetric observations should not shade the fact that heat flux and the baseline are smooth functions. Let us accept that they both belong to the $C^2(t_1,t_n)$ class of twice continuously differentiable functions. In this case, the discrete representation of heat flux can easily be converted to an analytical description by constructing an interpolating cubic spline $\dot{Q}(t)$ (Appendix explains how to calculate the spline). Although the spline can describe the whole array of experimental data points, only its portion corresponding to the thermal event (i.e. to the peak) is useful. Similarly, the baseline can be represented by an interpolating cubic spline $g(t)$. While constructing the baseline, one should keep in mind that the spline $g(t)$ must comply with the following conditions:

1 At the beginning of the transformation, the baseline coincides with the flux curve: $g(t_{beg})=\dot{Q}(t_{beg})$. Since no drastic deviations of the baseline from $\dot{Q}(t)$ are expected in the very beginning of the transformation, it is assumed that slopes (but not curvatures!) of the two splines are the same: $(dg/dt)_{t=t_{beg}}=(d\dot{Q}/dt)_{t=t_{beg}}$.
2 At the very end of the transition, the baseline smoothly approaches the flux curve: $g(t_{end})=\dot{Q}(t_{end})$, $(dg/dt)_{t=t_{end}}=(d\dot{Q}/dt)_{t=t_{beg}}$.
3 Expression (6) must be satisfied in all experimental points situated between t_{beg} and t_{end}. In other words, if t_{beg} corresponds to the point number i_{beg} and t_{end} corresponds to point number i_{end}, then (6) must be valid in $i_{end}-i_{beg}-1$ points between them:

$$g(t_i)=g(t_{beg})+$$

$$+\frac{g(t_{end})-g(t_{beg})}{\int_{t_{beg}}^{t_{end}}[\dot{Q}(\tau)-g(\tau)]d\tau}\int_{t_{beg}}^{t_i}[\dot{Q}(\tau)-g(\tau)]d\tau,$$

$$i=i_{beg}+1,...,i_{end}-1 \quad (7)$$

Since modern calorimeters are usually equipped with data acquisition systems capable of registering heat flux frequently (several times per second), expression (7), which is much easier to handle computationally, is a virtual twin of (6).

Based on the above conditions, an algorithm for constructing the spline $g(t)$ representing the integral baseline is presented below:

1 Choose an experimental point i_{beg} associated with the onset of transformation (if necessary, visualize calorimetric data). Estimate the slope of the flux curve at this point (employ, for instance, a simple linear regression for this purpose). Pick up an experimental point i_{end} corresponding to the end of transition and evaluate the slope at it. These first derivatives of the heat flux with respect to time are always used as boundary conditions needed for constructing splines. The first step of the algorithm is illustrated in Fig. 1. It should be accentuated that since a cubic spline is a local interpolant (Appendix), there is no need for very accurate estimations of slopes.
2 Use the experimental quantities t_i, $\dot{Q}_i$, $i_{beg},...,i_{end}$ along with the boundary conditions for constructing the spline $\dot{Q}(t)$. Once built, this spline remains intact for the rest of calculations. This step is exemplified in Fig. 2.

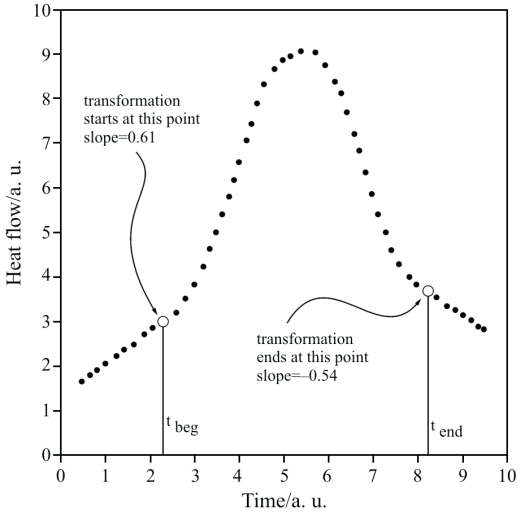

Fig. 1 Points corresponding to the onset and finish of a transformation and estimated slopes at these points

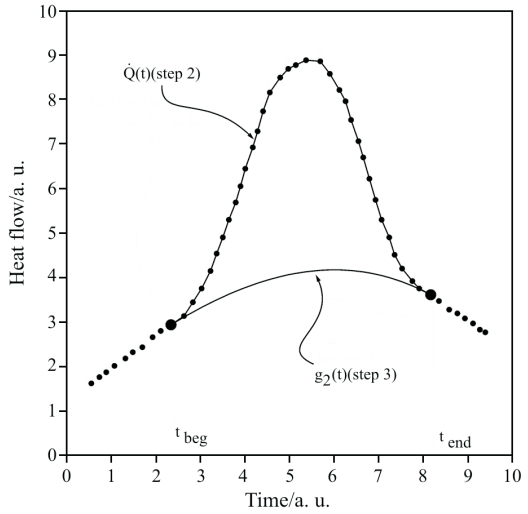

Fig. 2 Experimental points and the slopes are used for building a spline describing the heat flux curve and a spline $g_2(t)$ connecting the first and last points of the peak

3 Using the values of heat flux and its first derivatives at the knots $t_{i_{beg}}$ and $t_{i_{beg}}$ build the spline $g_2(t)$ ($g_2(t)$ is used instead of $g(t)$ for emphasizing that only two experimental measurements have been used for constructing the spline). Although $g_2(t)$ satisfies the first and second conditions formulated above ($g_2(t)$ and $\dot{Q}(t)$ have the same slopes in the first and last knots), it cannot be guaranteed that condition (7) is fulfilled. An example of the $g_2(t)$ spline is presented in Fig. 3. It is worth highlighting that a function identical to $g_2(t)$ was identified with the baseline by van der Plaats [3].

4 Having the analytical description $g_2(t)$, calculate $g_2(t_i)$, $i=i_{beg}+1,...,i_{end}-1$. After that, use the quantities t_i, $g_2(t_i)$, $i=i_{beg},...,i_{end}$ along with the fixed slopes in the first and last knots (these first derivatives, which

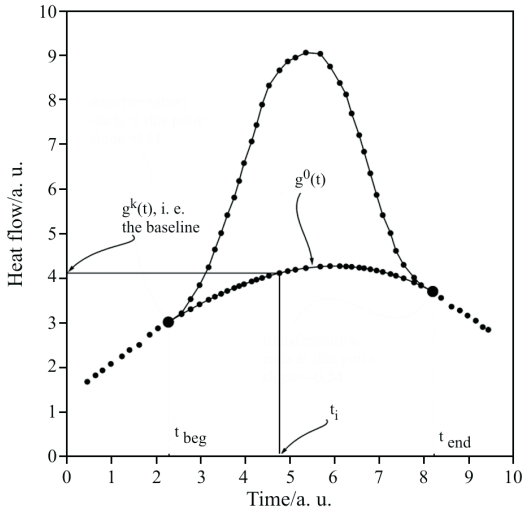

Fig. 3 Calculated $g_2(t)$ in all peak points are used for building an initial approximation to a baseline

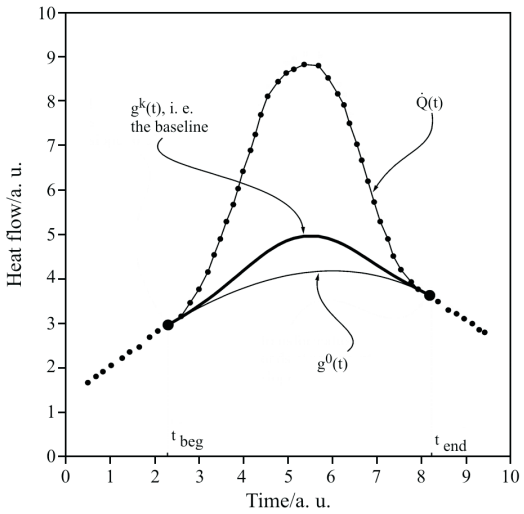

Fig. 4 Iterations resulting in a self-consistent baseline

are fixed once and forever, are always utilized for setting boundary conditions) for building the spline $g^0(t)$. The upper index k in $g^k(t)$ symbolizes the number of iteration undertaken so far for approaching the solution. While $g^0(t)$ is indistinguishable from $g_2(t)$, the former spline has the same number of knots as $\dot{Q}(t)$. The spline $g^0(t)$ is shown in Fig. 4.

5 Since both $\dot{Q}(t)$ and $g^0(t)$ are pieces of cubic polynomials smoothly glued together in the knots, taking integrals in (7) is trivial. Compute all $i_{end}-i_{beg}-1$ left-hand-sides of (7) and calculate corresponding entities in right-hand-sides. Use the latter values along with known and fixed boundary conditions for building $g^1(t)$.

6 Continue refining the spline by invoking (7) until the following condition is satisfied:

$$\frac{\left|g^k(t_i)-g^{k-1}(t_i)\right|}{1+\min\left[\left|g^k(t_i)\right|,\left|g^{k-1}(t_i)\right|\right]}\le\varepsilon,\quad \forall i=i_{beg}+1,\ldots,i_{end}-1 \tag{8}$$

Once (8) is respected, $g^k(t)$ is the baseline (depicted in Fig. 4) whose determination was the sole objective of all preceding computations. The criterion (8) recommended in [4] bears features of both absolute and relative accuracies. If the heat flux is small, (8) reduces to $\left|g^k(t_i)-g^{k-1}(t_i)\right|\le\varepsilon$, i.e. to a criterion based on absolute accuracy. If the heat flux is large, (8) becomes

$$\frac{\left|g^k(t_i)-g^{k-1}(t_i)\right|}{\min\left[\left|g^k(t_i)\right|,\left|g^{k-1}(t_i)\right|\right]}\le\varepsilon$$

which is a criterion based on relative accuracy.

Illustration of applicability

Since the approach has been tried numerous times, a collection of examples, from which a particular one is to be fished out, is large. The example presented here reflects authors' interests in applying calorimetry to an investigation of recrystallization and related phenomena in heat-treatable aluminum alloys [5, 6]. A piece of hot-rolled alloy was solutionized, quenched in water and deformed by cold rolling. A small sample was cut out from this piece. In the course of heating, a thermal event was registered by Perkin-Elmer Diamond differential power scanning calorimeter (pure annealed aluminum is used as a reference substance). It should be noted that an explicit usage of a baseline is deemed unnecessary for collecting qualitative information by DSC. Any obvious deviation from the baseline linearity is readily assigned to an exothermic or an endothermic effect [7]. Results of a microstructural investigation suggest that this event is likely precipitation of a new phase. A typical dependency of heat flux *vs.* time is shown in Fig. 5 along with two points corresponding to the beginning and end of transformation (a rigorous choice of such points can hardly be formalized). A dashed line, which smoothly originates from and smoothly comes back to the signal curve, is an initial approximation to the baseline, the approximation that does not satisfy the condition (6). Although ε in the right-hand-side of (8) is taken equal to 10^{-7}, only four iterations were needed for the procedure to converge. Since the spline obtained at the last iteration fulfills the condition that a degree of transformation is directly proportional to heat measured, it thus can be declared a self-consistent integral baseline. Heat flux with respect to the baseline and a time dependency of heat released are depicted in Fig. 6. Although the principles of the

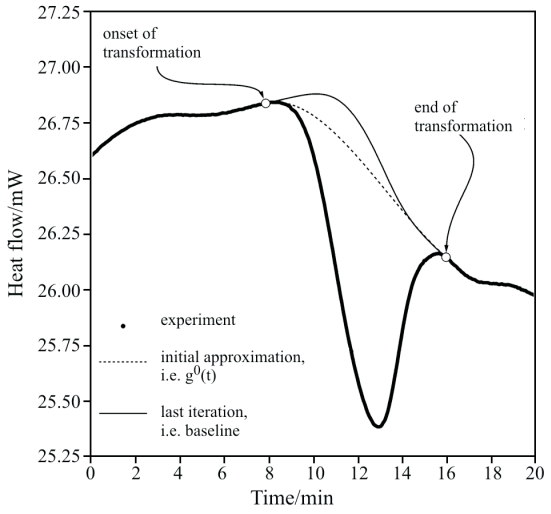

Fig. 5 An initial approximation and a self-consistent integral baseline constructed for an experimentally measured calorimetric response of a deformed Al alloy to heating

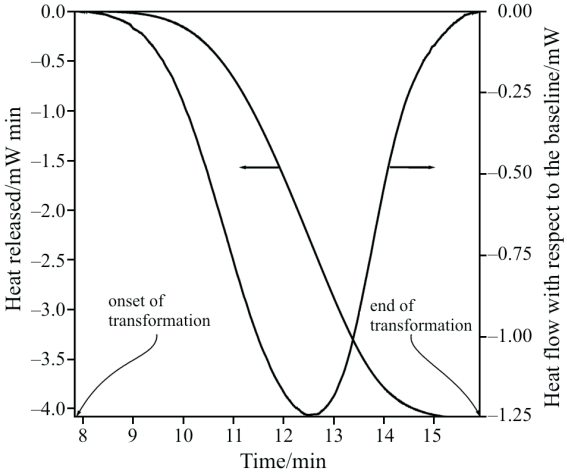

Fig. 6 Evolutions of heat flux and heat released with time

new approach were clarified in the preceding section by considering an endothermic transition, the example given clearly indicates that it is equally applicable to exothermic transformations: it is immaterial whether a peak is directed upward or downward.

Conclusions

The proposed method of constructing a baseline for the case when a degree of transformation is directly proportional to heat consumed or released has three distinctive features.

Firstly, it is assumed that both a directly measured calorimetric signal and a corresponding baseline are smooth functions of time. In addition to this assumption, which is not difficult for the calorimetric community to accept, no further simplifications are employed. Moreover, the approach is free from such 'numerically fragile' undertakings as extrapolating a polynomial description of heat capacity into the peak range [3, 8].

Secondly, the iterative procedure described above leads to the integral baseline satisfying condition (6). This circumstance along with the fact that the baseline is indistinguishable from a heat flux curve before and after the transformation justifies the term 'self-consistent' in the title of this contribution.

Thirdly, the suggested method of constructing the integral baseline is very simple to apply in practice. This straightforwardness stems from utilizing interpolating cubic splines for representing both the peak and baseline. In contrast to a habitual interpolating polynomial, a spline is a local interpolant. Thus it connects experimental points without having aphysical behavior between them. If experimental data are noisy, then approximating (smoothing) splines can be considered, but a necessity to choose a smoothing parameter makes the approach less universal and more computationally cumbersome. It is worth mentioning that splines de-

scribing the peak and baseline are never differentiated, they are only integrated, which ensures a computational robustness of the method. It seems that despite of a number of publications devoted to building the integral baseline (a great deal of them is cited in [9]), a usefulness of splines for this purpose has been overlooked.

It is worth mentioning that the sensitivity and accuracy of modern calorimeters are tried to be improved through eliminating noises and drifts in the baseline [10]. These 'hardware efforts' are to be supported by 'software efforts', indeed. It is hoped that the technique proposed will be utilized in data acquisition and analysis software packages accompanying modern calorimeters.

Appendix

A cubic spline is a function defined on a mesh of knots $x_1,...,x_n$, $x_i<x_{i+1}$, $i=1,...,n-1$. Within each interval $[x_i,x_{i+1}]$, the spline is a piece of a cubic polynomial. Adjacent polynomials are 'glued' in the internal knots in such a way that the spline is twice continuously differentiable function on $[x_1,x_n]$. Ready-to-use subroutines for building an interpolating cubic spline can be found in many publicly available libraries including IMSL** and HSL***. An elegant and easily implementable algorithm presented below is intended for readers who decide to write their own program for constructing the integral baseline instead of relying upon a black box. A FORTRAN 77 source code implementing the algorithm is available for free from the authors upon request.

The second derivative of the cubic spline is a linear function on $[x_i,x_{i+1}]$:

$$Sp''(x) = m_{i+1}\frac{x-x_i}{h_i} + m_i\frac{x_{i+1}-x}{h_i},\ i=1,...,n-1 \qquad (A1)$$

where m_i and m_{i+1} are the second derivatives in knots x_i and x_{i+1}, correspondingly, and $h_i \equiv x_{i+1}-x_i$. The expression (A1) ensures the continuity of the second derivative in the inner knots $x_2,...,x_{n-1}$. To obtain the expression for the spline, (A1) has to be integrated twice:

$$Sp(x) = \frac{m_{i+1}}{6h_i}(x-x_i)^3 + \frac{m_i}{6h_i}(x_{i+1}-x)^3 + \frac{C_{i+1}}{h_i}(x-x_i) + \frac{C_i}{h_i}(x_{i+1}-x) \qquad (A2)$$

where C_i and C_{i+1} are integration constants. They can be determined using known values of function in knots x_i and x_{i+1}:

$$Sp(x_i) = f_i = \frac{m_i h_i^2}{6} + C_i \Rightarrow C_i = f_i - \frac{m_i h_i^2}{6} \qquad (A3)$$

$$Sp(x_{i+1}) = f_{i+1} = \frac{m_{i+1}h_i^2}{6} + C_{i+1} \Rightarrow C_{i+1} = f_{i+1} - \frac{m_{i+1}h_i^2}{6} \qquad (A4)$$

Substituting (A3) and (A4) into (A2) and then differentiating the expression obtained, one arrives at the following representation of the first derivative on $[x_i,x_{i+1}]$:

$$Sp'(x) = \frac{m_{i+1}}{2h_i}(x-x_i)^2 - \frac{m_i}{2h_i}(x_i+1-x)^2 + \left(\frac{f_{i+1}}{h_i} - \frac{m_{i+1}h_i}{6}\right) - \left(\frac{f_i}{h_i} - \frac{m_i h_i}{6}\right)$$

The requirement that the first derivative of spline must be continuous in the inner knots leads to the following system of linear equations with respect to second derivatives:

$$m_{i-1}\frac{h_{i-1}}{6} + m_i\left(\frac{h_{i-1}}{3} + \frac{h_i}{3}\right) + m_{i+1}\frac{h_i}{6} = \frac{f_{i-1}}{h_{i-1}} - f_i\left(\frac{1}{h_{i-1}} + \frac{1}{h_i}\right) + \frac{f_i}{h_i}$$
$$i=2,...,n-1 \qquad (A5)$$

While the total number of unknowns is equal to n, the system (A5) contains only $n-2$ equations. If an accurate estimation of the first or second derivative is available in the first and last knots, they can be used as such conditions. Otherwise, the so-called 'not-a-knot' conditions are employed:

$$Sp''(x_2-0) = Sp''(x_2+0) \Rightarrow m_1 - 2m_2 + m_3 = 0$$
$$Sp''(x_{n-1}-0) = Sp''(x_{n-1}+0) \Rightarrow m_{n-2} - 2m_{n-1} + m_n = 0 \qquad (A6)$$

Whatever the choice, two equations representing two boundary conditions can be added to the system (A5) in such a way that its matrix will be tridiagonal with diagonal dominance. The latter means that it is positive definite, which, in turn, prompts that an extremely simple, fast and robust algorithm can be used for building the interpolating cubic spline.

References

1 J. Dupuy, E. Leroy and A. Maazouz, J. Appl. Polym. Sci., 78 (2000) 2262.
2 W. Hemminger and G. Höhne, Calorimetry: Fundamentals and Practice, Verlag Chemie, Deerfield Beach 1984, p. 310.
3 G. van der Plaats, Thermochim. Acta, 72 (1984) 77.
4 P. E. Gill, W. Murray and M. H. Wrigh, Practical optimization, Academic Press, London, New York 1981, p. 401.
5 W. H. Hildebrandt, Metall. Trans. A, 10A (1979) 1045.
6 F. Haessner, Calorimetric Investigationo Recovery and Recrystallization in Metals, Thermal Analysis in Metallurgy, Ed. R. D. Shull, A. Joshi, TMS, Warrendale 1992, p. 233.
7 G. Riontino, G. L. Ferro, D. Lussana and M. Massazza, J. Therm. Anal. Cal., 82 (2005) 83.
8 W. F. Hemminger and S. M. Sarge, J. Thermal Anal., 37 (1991) 1455.
9 G. W. H. Höhne, W. Hemminger and H. J. Flammersheim, Differential Scanning Calorimetry, Springer, Berlin, New York 2003, p. 298.
10 S. Wang, K. Tozaki, H. Hayashi and H. Inaba, J. Therm. Anal. Cal., 79 (2005) 605.

Received: May 29, 2006
Accepted: October 17, 2006

DOI: 10.1007/s10973-006-7702-3

** http://www.visualnumerics.com/products/imsl/index.html
*** http://www.aspentech.com/hsl/default.asp