Decomposition of the vibrational heat capacity

1 - Theoretical description of the analysis mechanism

[Heat capacity] is a fundamental physical quantity, widely used in studying of solids, especially in case of the presence of phase transitions (PT). Today, commercial calorimeters have become available for accurate measurements, producing high quality data that afford insightful analysis, giving us crucial knowledge about the mechanism of PT. Heat capacity also relates to [critical exponent] α , describing its behavior near PT.

Without PT heat capacity (called **"vibrational heat capacity"**) could be represented by harmonic oscillator models, such as [Debye] and [Einstein] functions. For the heat capacity C_V at constant volume one can use a combination of these functions to decompose its dependence $C_V(T)$ from temperature T. For example, a diatomic ionic crystal NaCl (sodium chloride) comprises [one Debye and one Einstein function].

In most studies, however, a temperature dependence of heat capacity at constant pressure $C_p(T)$ is measured. Furthermore, the anomalous part of it (due to PT) needs to be separated from the vibrational one $C_p^{(vibrational)}(T)$ in order to analyze PT (Fig. 1). This dependence $\Delta C_p = C_p - C_p^{(vibrational)}$ is also called **"excess heat capacity"**.
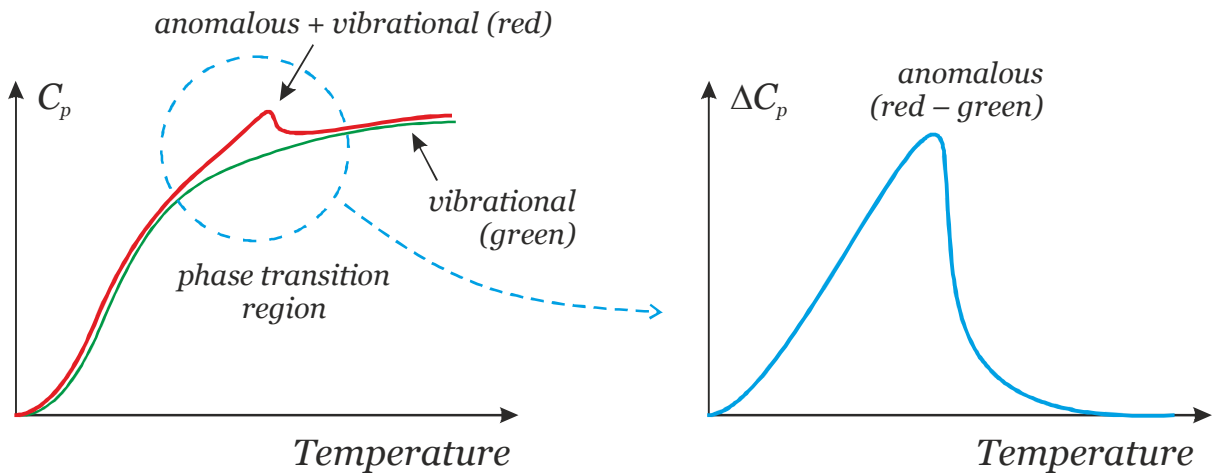

Fig. 1 – Heat capacity and its anomalous part (excess heat capacity)

Hence the algorithm is obvious: to obtain $\Delta C_p(T)$ (blue line in Fig. 1) one has to evaluate vibrational heat capacity $C_p^{(vibrational)}(T)$ (green line) and

subtract it from the experimentally measured $C_p(T)$ (red line), which contains the region of PT (dashed blue circle).

The most complicated step of this algorithm is to decompose vibrational (sometimes also called **"background"**) part analytically. For convenience, let's use simpler notation here and below as $C_p^{(vibrational)}(T) \equiv C_{bg}(T)$. It remains still a heat capacity _at constant pressure_, though.

First of all, by excluding experimental points (observations) from $C_p(T)$ in PT region (dashed blue circle), we can obtain a subset of data as if there is no PT at all (in other words, _experimental_ set for background $C_{bg}^{(exp)}(T)$). Next, by using a proper nonlinear model, we could fit a regression curve (green on Fig. 1) $C_{bg}(T)$.

The main reference, used in this work, is next:

[**T. Matsuo, Thermochimica Acta 267, 421 (1995)**]

According to the authors, background C_{bg} could be described as:

$$C_{bg} = C_{internal} + C_{external} + \left(C_p - C_V\right). \tag{1}$$

Here $C_{internal}$ term corresponds to the heat capacity of intramolecular vibrations, while $C_{external}$ represents heat capacity due to lattice vibrations (in single crystals). The model includes also a small correction term for the difference between C_p and C_V, so we can use combination of Einstein (C_E) and Debye (C_D) models, as mentioned above.

Both Einstein and Debye theories are based on harmonic oscillator models, but characterized temperatures are used as parameters instead of frequencies (Einstein and Debye temperatures Θ_E and Θ_D, respectively):

$$\begin{aligned} C_E\left(\Theta_E, T\right) &= \frac{\left(\Theta_E/T\right)^2 \exp\left(-\Theta_E/T\right)}{\left(1 - \exp\left(-\Theta_E/T\right)\right)^2}, \\ C_D\left(\Theta_D, T\right) &= 3\left(T/\Theta_D\right)^3 \int_0^{\Theta_D/T} \frac{x^4 \exp(x)}{\left(\exp(x) - 1\right)^2} dx. \end{aligned} \tag{2}$$

For a particular solid with _N_ atoms in a cell (molecule in lattice), there is 3_N_ vibrational modes. According to the structure of the lattice, one can evaluate number of internal vibrations and external. **It is known**, that the internal component $C_{internal}$ comprises **only Einstein functions** (temperatures Θ_E can be found _in the spectroscopy_), whilst the external $C_{external}$ – **Einstein and Debye** (_generally unknown_ and matched to the experimental data). Each term contain

different number of functions – $N_E(1)$ for the $C_{internal}$, $N_E(2)$ and $N_D(2)$ for $C_{external}$, respectively:

$$C_{internal} = R \sum_{i=1}^{N_E(1)} g_{Ei}(1) C_E\left(\Theta_{Ei}(1), T\right)$$

$$C_{external} = R\left[\sum_{i=1}^{N_E(2)} g_{Ei}(2) C_E\left(\Theta_{Ei}(2), T\right) + \sum_{i=1}^{N_D(2)} g_{Di}(2) C_D\left(\Theta_{Di}(2), T\right) \right] \tag{3}$$

$$3N = \sum_{i=1}^{N_E(1)} g_{Ei}(1) + \sum_{i=1}^{N_E(2)} g_{Ei}(2) + \sum_{i=1}^{N_D(1)} g_{Di}(2).$$

Here $g_{E,D}$ are the corresponding weights of each component, R is a <u>gas constant</u>.

Finally, a correction term can be set as:

$$\left(C_p - C_V\right) \approx ATC_{external}^2, \tag{4}$$

where A is an additional fitting parameter. All terms from the model are visualized in Fig. 2.
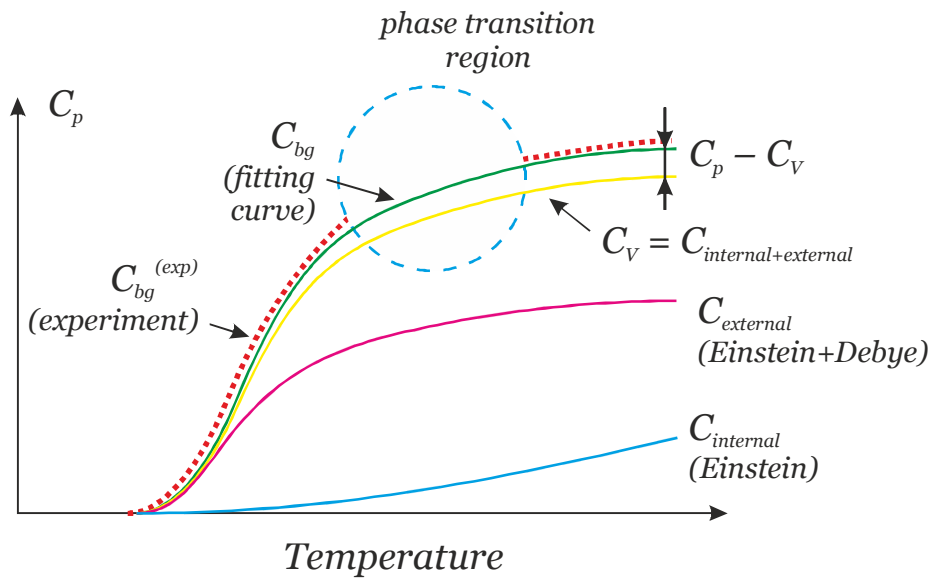

Fig. 2 – Decomposition of the vibrational heat capacity

<u>The main objective of this assignment</u> is to obtain the $C_{bg}(T)$ curve and use it to calculate the excess heat capacity $\Delta C_p(T)$. It requires collecting set of temperatures $\Theta_{Ei}(1)$ from the spectroscopy to compose $C_{internal}$, choosing the adequate regression model for the $C_{external}$ and fitting the experimental curve $C_{bg}^{(exp)}(T)$ in order to obtain the set of parameters $\Theta_{Ei}(2)$, $\Theta_{Di}(2)$ and A.

2 – The objects and input data

The objects used in this analysis are two solids with 2nd order PT: tri-rubidium deuterium disulfate **Rb$_3$D(SO$_4$)$_2$** (TRDS, PT at $T_C \approx 78.5K$) and lead germanate **Pb$_5$Ge$_3$O$_{11}$** (PGO, PT at $T_C \approx 450K$) single crystals. The temperature dependencies for the heat capacity in both compounds are of similar character to that shown in Fig. 1.

We use data from the next references:

1) The experimental data $C_p(T)$ of TRDS, temperatures $\Theta_{Ei}(1)$ with their weights $g_{Ei}(1)$ to compose $C_{internal}$ were taken from

[**T. Matsuo et al., Thermochimica Acta, 403, 137 (2003)**]

2) Spectroscopy data for $C_{internal}$ were also analyzed from

[**N. Fourati et al., Phase Transitions, 18, 87 (1989)**]

3) The experimental data $C_p(T)$ of PGO were composed, using data from

[**A. A. Bush and E. A. Popova, Fiz. Tverd. Tela 46, 875 (2004)**]

[**S. A. Ivanov et al., Fiz. Tverd. Tela 21, 2545 (1979)**]

4) Spectroscopy data for $C_{internal}$ were taken from

[**D. J. Lockwood et al., J. Phys. C: Solid St. Phys., 13, 1539 (1980) 1539**]

3 – Internal vibrations

There are $N = 14$ atoms in the TRDS chemical unit and therefore $3N = 42$ vibrational modes (18 – internal, 24 – external). We replicated $C_{internal}$ from (1) ($\Theta_{Ei}(1)$ temperatures with their weights $g_{Ei}(1)$) as proposed by the authors:

Rb$_3$D(SO$_4$)$_2$	
Einstein temperatures $\Theta_{Ei}(1)$, K	Weights $g_{Ei}(1)$
1451	2
635	4
1683	6
879	8

For the PGO crystal there are $N = 19$ atoms in the chemical unit, hence $3N = 57$ vibrational modes (30 – internal (9 of which relate to tetrahedrons GeO$_4$$^{4-}$, 21 – to double tetrahedrons Ge$_2$O$_7$$^{6-}$), 27 - external). We have

composed our own set using Raman spectrum data (D. J. Lockwood et al.) by providing exploratory analysis:

Pb$_5$Ge$_3$O$_{11}$			
(Ge$_2$O$_7$$^{6-}$):		(GeO$_4$$^{4-}$):	
Einstein temperatures $\Theta_{Ei}(1)$, K	Weights $g_{Ei}(1)$	Einstein temperatures $\Theta_{Ei}(1)$, K	Weights $g_{Ei}(1)$
444	2	589	1
453	2	639	1
479	4	712	1
499	3	1009	2
507	3	1182	1
536	6	1136	1
570	1	1161	2

Data from tables above represent $C_{internal}$ term from (1) that corresponds to internal vibrations.

4 – External vibrations (fitting procedure)

4.1 – Choosing set of Einstein and Debye terms

To compose $C_{external}$ term (see eqs. (1) and (3)), which corresponds to external vibrations, one has to choose a reasonable set of temperatures $\Theta_{Ei}(2)$, $\Theta_{Di}(2)$ with their weights $g_{Ei}(2)$ and $g_{Di}(2)$. Too many of them may result in complicated regression model and fitting problems with high correlations between parameters.

The nonlinear model (1), used in fitting, is already complicated by Einstein and (especially) Debye functions. To build a flexible code, we use **Lmfit**, a special extension package for Python – a high-level interface to non-linear optimization and curve fitting problems. It allows computing our problem by minimizing residual (i.e. data-model) array $\left\{ C_{bg}^{(exp)}(T) - C_{bg}(T) \right\}$.

More information could be found in <u>lmfit github repository</u>.

There are several methods to verify if the data is consistent with the model or not. One of them proposes repeatable fitting by the same model (1), using subsets from the experimental data $C_{bg}^{(exp)}(T)$ for different temperature ranges – from narrow to full span. Thus one can obtain a dependence of fitting

parameters ($\Theta_{Ei}(2)$ and $\Theta_{Di}(2)$) from the upper limit temperature of the interval used for the fitting $\Theta(T_u)$, or from the number of points N_T, used for fitting $\Theta(N_T)$. The good model means similar results for $\Theta_{Ei}(2)$ and $\Theta_{Di}(2)$ for different fittings, with little changes in $\Theta(N_T)$ curves (ideally constants).

To provide such analysis, we wrote the next Python scripts:

TRDS_3temperatures.py (proposed three-temperature model for TRDS)
TRDS_4temperatures.py (original four-temperature model for TRDS)
PGO_temperatures.py (model for PGO)

The scripts are run from the command line as:

PGO_temperatures.py <data file name> <output file name>
TRDS_4temperatures.py <data file name> <output file name>
TRDS_4temperatures.py <data file name> <output file name>

Here **<data file name>** is an input data file with $C_{bg}^{(exp)}(T)$, "background" heat capacity subset (without observations that include phase transition). It contains two columns – temperature and heat capacity. For the TRDS, the input file is *TRDS_bg.dat* (normalized by gas constant $C_{bg}^{(exp)}/R(T)$), for the PGO – *PGO_bg.dat*. The output file **<output file name>** will contain table of $\Theta(N_T)$ dependencies (*TRDS_3temps.dat, TRDS_4temps.dat* and *PGO_temps.dat*).

In case of TRDS, from the output (see .dat-files above, also with pictures *TRDS_4temperatures_N_points.png* and *TRDS_3temperatures_N_points.png*) it is clearly visible, that the regression model for $C_{external}$ with 4 temperatures (3 Einstein and 1 Debye), used by the authors is not the best one, because two of the temperatures $\Theta_{Ei}(2)$ are nearly the same. We proposed an alternative 3 temperatures model (**TRDS_3temperatures.py**) with **2 Einstein functions and 1 Debye with the weights 11, 4 and 9, respectively**. For the PGO, we have created a model with **1 Einstein and 2 Debye functions (weights 8, 12 and 7, respectively)**. It is quite reasonable set, as seen from the output picture *PGO_temperatures_N_points.png*.

4.2 – The main scripts

The main scripts for this assignment are **PGO_fit.py** and **TRDS_fit.py**. They estimate set of fitting parameters $\Theta_{Ei}(2)$, $\Theta_{Di}(2)$ and $\left(C_p - C_V\right)$ correction term A for each solids, and build the excess heat capacity $\Delta C_p(T)$,

showing results in graphs and writing them in output files. From the command line they run as:

PGO_fit.py <**background input**> <**full data input**> <**output file**>
TRDS_fit.py <**background input**> <**full data input**> <**output file**>

Here <**background input**> is an input data file with $C_{bg}^{(exp)}(T)$ (in our case *PGO_bg.dat* and *TRDS_bg.dat*), <**full data input**> is a data file with all experimental points $C_p(T)$, including PT region (files *PGO_fullPT.dat* and *TRDS_fullPT.dat*), and <**output file**> is a data file with a table of 3 columns: temperature, decomposed background and excess heat capacity (stored in files *PGO_excess.dat* and *TRDS_excess.dat*).

4.3 – Estimating errors

A common problem in nonlinear regression is estimating errors. Asymptotic Standard Errors (ASE) *should not be used for nonlinear models*, since they underestimate the magnitude of the parameter uncertainties. The ASE are based on the "information matrix" and they ignore the off-diagonal elements. Probably the best alternative is using the **Bootstrap approach**. It provides multiple curve-fitting, seeding the data with randomly sampled residuals $\left\{ C_{bg}^{(exp)}(T) - C_{bg}(T) \right\}$. Histograms of the parameters are obtained as the result.

Finally, mathematical model $C_{bg}(T)$ could be verified by performing the **Runs Test**. It does not validate the model itself, it simply says if the data and/or assumptions about the data are not consistent with the model. As the output Runs test evaluates standard score (also *Z*-score) to verify consistency.

More on this: **doi:10.1016/S0091-679X(07)84024-6**

Z-scores, as seen from the output, printed in *stdout.txt* for both solids, are extremely high (-9.225 and -11.417). They point at more complex set of factors that influence heat capacity, including PT itself.