**Chemometric methods for Raman spectrum calibration

Bastian Barton, PhD

Fraunhofer LBF, Division Plastics, Materials Analysis Group

Schlossgartenstr. 6, D-64289 Darmstadt, Germany

# Wavenumber calibration

## Requirements:

* Standard material (SM): The SM should generate a Raman spectrum with a large number of well-separated, equally spaced Raman peaks over a large wavenumber range. The requirements for SM are defined in WP3.
* A list of absolute anchor point (e.g. peak) positions, or a standard Raman spectrum (both to be provided by CHARISMA)

## Wavenumber calibration procedure for user instrument:

1. A calibration spectrum (CS) of the standard material is measured at the user site. The measurement should be carried out frequently, and under “typical” operating conditions (temperature, humidity, laser power etc.).
2. A defined set of anchor points (AP) is located on the CS (Figure 1), e.g. by

* Peak maxima (after peak fit by Gauss / Lorentz)
* Centers of mass of peaks

1. For each AP, the shift list is calculated: *si* = *APSM,i*  – *APCS,i*
2. The *si* are plotted in a graph against the CS wave vector: *si(kCS)*
3. A continuous, parametrized curve *s-1*(*k*) is fitted through the *si(kCS)*, e.g. a polynomial or spline (Figure 2b). This wavenumber calibration curve is stored and assigned to the specific user instrument, and given a timestamp.

## Transformation of spectra for CHADA

For CHADA generation, any measured Raman spectrum is transformed according to the current wavenumber calibration of the recording instrument. For that, the wavenumber vector *k* from the measurement is transformed as *k’* = *k* + *s-1*(*k*) (Figure 2d).

This will result in a nun-uniform wavenumber sampling, since the *k’* will no longer be equally spaced. Therefore, the wavenumber calibration should be performed before interpolation of the spectrum to a different wavenumber axis, and before amplitude calibration.

Alternatively, the transformer (containing *s-1*(*k*) and the timestamp) can be stored with the data, and the transformation carried out at a later point in time - e.g. for classification, data base search, comparison or multivariate processing within CHADA.

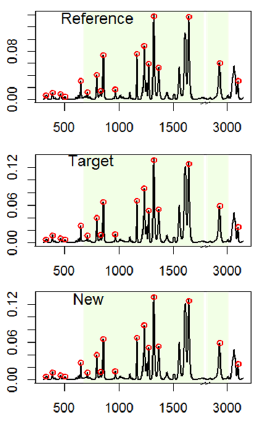


Figure 1: Anchor points (maxima, red) in reference spectrum. The AP of the reference may be given as a list. From [1].

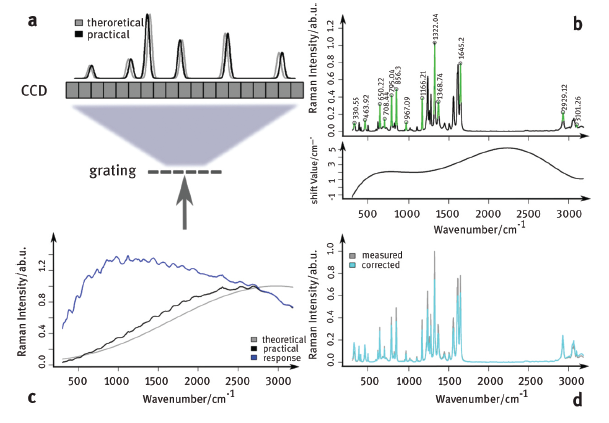


Figure 2: Workflow of wavenumber (a–b) and intensity calibration (c–d). (a) The relation between wavenumber and pixel

positions can change, leading to wavenumber misalignment between measured and theoretical Raman spectra. (b) The

wavenumber misalignment is corrected based on Raman spectra of a known standard material. (c) The intensity response

function of the device is calculated as the ratio between measured and theoretical emission of a known standard material.

(d) Intensity axis of measured Raman spectra is corrected by the calculated intensity response function. From [2].

## Python pseudocode

T\_wn = chada.wavelengthTransformer()

T\_wn.fit(calibration\_spec, list\_of\_standard\_peak\_pos)

chada\_spec = T\_wn.transform(measured\_spec)

# Amplitude calibration (gain correction)

## Requirements:

* Standard material (SM): The SM should generate counts > 0 within a large wavenumber range (e.g. SRM 2242, National Institute of Standards & Technology, Gaithersburg, MD 20899, USA). Defined peaks are not reuqired. The requirements for SM are defined in WP3.
* Raman spectrum of the SM (SMS) measured by a “standardized” (calibrated) system, or theoretical Raman spectrum

## Amplitude (gain) calibration procedure for user instrument:

1. A calibration spectrum (CS) of the standard material is measured at the user site. The measurement should be carried out frequently, and under “typical” operating conditions (temperature, humidity, laser power etc.).
2. The wavenumber vectors of the CS are calibrated as described under 1.
3. The gain curve *g* is calculated: *g*(*k*) = *ySMS*(*k*) / *yCS*(*k*). This gain curve is stored and assigned to the specific user instrument, and given a timestamp.

## Transformation of spectra for CHADA

For CHADA generation, any measured Raman spectrum is transformed according to the current gain curve of the recording instrument. For that, the count vector *y* from the measurement is transformed as *y’(k)* = *g*(*k*) *y(k)* (Figure 2c).

Alternatively, the transformer (containing *g*(*k*) and the timestamp) can be stored with the data, and the transformation carried out at a later point in time - e.g. for classification, data base search, comparison or multivariate processing within CHADA.

## Python pseudocode

T\_amp = chada.amplitudeTransformer()

T\_amp.fit(calibration\_spec, standard\_spec)

chada\_spec = T\_amp.transform(measured\_spec)

# Potential further calibrations

In principle, transformers can be fitted for other instrument-dependent artifacts being a scalar function of the wavenumber k, e.g.

* dark current
* peak broadening

These transformers can be fitted by calibration with appropriate standards, and can be added to the CHADA transformers list as needed.

# Literature

[1] Guo, S., Heinke, R., Stöckel, S., Rösch, P., Bocklitz, T., & Popp, J. (2017). Towards an improvement of model transferability for Raman spectroscopy in biological applications. Vibrational Spectroscopy, 91, 111-118.

[2] Ryabchykov, O., Guo, S., & Bocklitz, T. (2018). Analyzing Raman spectroscopic data. Physical Sciences Reviews, 4(2).