Supplementary Materials

The perfection of Raman spectroscopic gas densimeters

Ronald J. Bakker

Resource Mineralogy, Department of Applied Geosciences and Geophysics
Montanuniversity Leoben, Peter-Tunner Str. 5, 8700 Leoben, Austria

Email: bakker@unileoben.ac.at

S1 | Uncertainty

The uncertainty is a fundamental property of scientific data values, and it expresses a state of limited knowledge where it is impossible to exactly describe the existing state (e.g. Drosg, 2009). Uncertainties can be estimated for a single measurement and for multiple measurements of the same object. The present work uses the uncertainty of single measurements, which is limited by the precision of the spectrometer, along with other factors that might affect the ability to make the measurement. Calculations with these uncertainties are performed with the best and worse-case scenarios (upper-lower bounds method).

Multiple measurements can be analysed with probability distribution functions. The uncertainty may be defined by the range around the best estimate with a 100% confidence (maximum uncertainty). The best estimate can be defined as an average value, and the standard deviation (standard uncertainty) represents the 1σ confidence interval (i.e. about 68.27% for a Gaussian distribution). Most publications about the reliability of Raman band peak position estimations include this statistical method to define uncertainties, and do not consider precision of the spectrometer. In addition, probability distribution functions are also used in a completely different context for Raman spectra, i.e. mainly to reproduce the morphological properties of Raman bands. For example, the peak position of a Raman band is 50 times estimated with a specific best-fit probability distribution function, and these peak values are submitted to a statistical analysis, resulting in a best estimate (average) and a calculated uncertainty (confidence interval).

Reference

Drosg M (2009) *Dealing with uncertainties, a guide to error analysis*, Springer, Heidelberg.

S2 | Probability Distribution Functions in the software LabSpec 5, 6 and PeakFit

S2.1 | Gaussian function (normal distribution)

Software LabSpec defines the parameters *p* (peak position), *a* (peak amplitude), and *w* (full width at half maximum, *FWHM*)

$$y_{Gauss} = a \cdot exp\left[-4ln2 \cdot \left(\frac{x-p}{q}\right)^2\right] \quad (1)$$

where *a* is maximum amplitude, *p* is centre position, *q* = *w* (= *FWHM*)

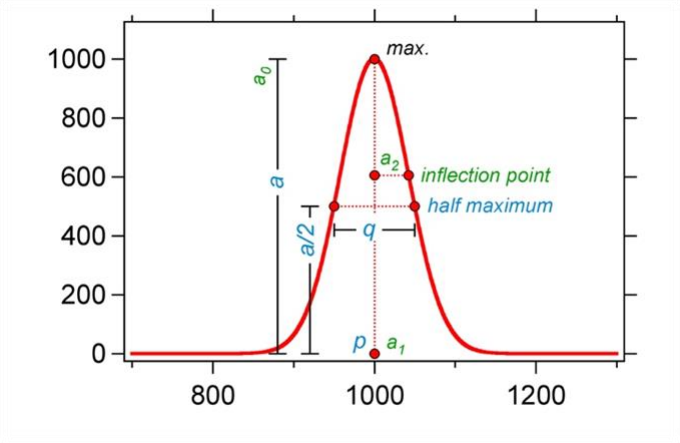

Figure S1. Example of a Gaussian function with geometrical definitions of the parameters (*a* = 1000, *p* = 1000, *q* = 100, *w* = 100, *FWHM* = 100).

Alternative and equivalent definition (e.g. PeakFit v4.12, SYSAT Software Inc.):

$$y_{Gauss} = a_0 exp\left[-\frac{1}{2}\left(\frac{x-a_1}{a_2}\right)^2\right] \quad (2)$$

where a_0 is maximum amplitude, a_1 is centre position, and a_2 is distance between inflection point and centre position (i.e. width).

$$FWHM = 2\sqrt{2ln2} \cdot a_2 \quad (3)$$

S2.2 | Lorentzian function (Cauchy distribution)

Software LabSpec defines the parameters *p* (peak position), *a* (peak amplitude), and *w* (full width at half maximum, *FWHM*)

$$y_{Lorentz} = \frac{a}{1+4\left(\frac{x-p}{q}\right)^2} \quad (4)$$

where *a* is maximum amplitude, *p* is centre position, *q* = *w* (= *FWHM*)

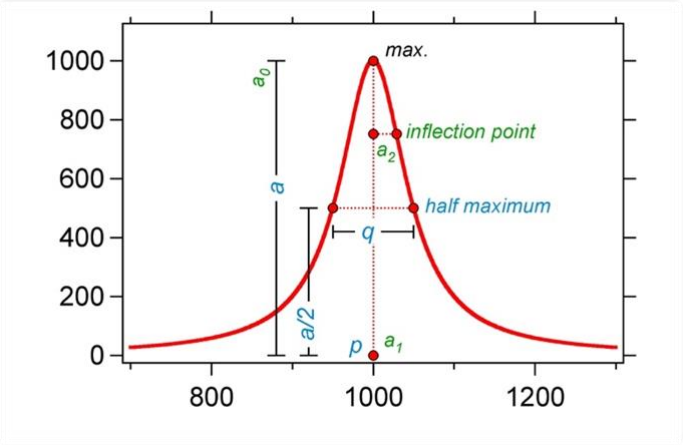

Figure S2. Example of a Gaussian function with geometrical definitions of the parameters (*a* = 1000, *p* = 1000, *q* = 100, *w* = 100, *FWHM* = 100).

Alternative and equivalent definition (e.g. PeakFit v4.12, SYSAT Software Inc.):

$$y_{Lorentz} = \frac{a_0}{1+\frac{4}{3}\left(\frac{x-a_1}{2a_2}\right)^2} \quad (5)$$

where a_0 is maximum amplitude, a_1 is centre position, and a_2 is distance between inflection point and centre position (i.e. width).

$$a_2 = \frac{1}{6}\sqrt{3} \cdot FWHM \quad (6)$$

S2.3 | Combination Gaussian-Lorentzian function (symmetric)

Software LabSpec defines the parameters *p* (peak position), *a* (peak amplitude), and *w* (full width at half maximum, *FWHM*) and *g* (Gaussian contribution)

LabSpec reference manuals do not provide any definition of mixed Gaussian and Lorentzian functions. A simple summation of fractions of individual distribution functions that have the same a, p, and w values can be deduced from manual analyses of spectra.

$$y_{GL} = g \cdot y_{Gauss} + (1 - g) \cdot y_{Lorentz} \tag{7}$$

the parameter *g* is the Gaussian fraction in the combined function (value between 0 and 1)

Alternative, but not equivalent definition (e.g. PeakFit v4.12, SYSAT Software Inc.):

$$y_{GL} = a_0 \left[\frac{\frac{a_3\sqrt{ln2}}{a_2\sqrt{\pi}} exp\left(-4ln2\left(\frac{x-a_1}{a_2}\right)^2\right) + \frac{1-a_3}{\pi a_2\left[1+4\left(\frac{x-a_1}{a_2}\right)^2\right]}}{\frac{a_3\sqrt{ln2}}{a_2\sqrt{\pi}} + \frac{1-a_3}{\pi a_2}} \right] \tag{8}$$

where a_0 is maximum amplitude, a_1 is centre position, a_2 is *FWHM*, and a_3 is the fraction of Gaussian (value between 0 and 1). Note that the definition of the a_2 parameter is different from the a_2 in a single Gaussian and a single Lorentzian function.

This function can be summarized as:

$$y_{GL} = \frac{\frac{a_3\sqrt{ln2}}{a_2\sqrt{\pi}} y_{Gauss} + \frac{1-a_3}{\pi a_2} y_{Lorentz}}{\frac{a_3\sqrt{ln2}}{a_2\sqrt{\pi}} + \frac{1-a_3}{\pi a_2}} \tag{9}$$

Comparison of both functions (7 and 8):

a distribution curve constructed with a_3 = 0.5 (half Gaussian and half Lorentzian) according to PeakFit v4.12 corresponds to *g* = 0.596 in the LabSpec 5 and 6 software.

S2.4 | Asymmetric Gaussian-Lorentzian function (LabSpec 6)

Software LabSpec defines the parameters *p* (peak position), *a* (peak amplitude), *w* (peak full width at half maximum, *FWHM*), *g* (Gaussian contribution), and *t* (asymmetry factor)

$$y_{aGL} = g \cdot y_{aGauss} + (1 - g) \cdot y_{aLorentz} \quad (10)$$

LabSpec reference manuals do not provide any definition of asymmetric mixed Gaussian and Lorentzian functions. Detailed analyses of the peak fitting application reveal that the symmetric functions are split in two halves (left and right of the centre position at *p*). Each half has the same maximum amplitude (*a*), but different *FWHM*.

In symmetric functions:

$$w = q = FWHM \quad (11)$$

$$\gamma = \tfrac{1}{2} \cdot q = \tfrac{1}{2} \cdot FWHM \quad (12)$$

New definition in asymmetric functions:

right part ($x > p$)

$$\gamma_{right} = \tfrac{1}{2} \cdot q_{right} \quad (13)$$

left part ($x < p$)

$$\gamma_{left} = t \cdot q_{right} \quad (14)$$

Real peak width at half maximum:

$$FWHM = \gamma_{left} + \gamma_{right} = \left(\tfrac{1}{2} + t\right) \cdot q_{right} \quad (15)$$

Symmetric functions are defined with *t* = 0.5

It should be noted that the parameter *w* is not equal to the *FWHM*, a mistake in the LabSpec 6 reference manuals:

$$w = q_{right} = \frac{FWHM}{\frac{1}{2}+t} \quad (16)$$

S2.5 | Asymmetric Gaussian function (LabSpec 6)

Right part: $$y_{aGauss} = a \cdot exp\left[-4ln2 \cdot \left(\frac{x-p}{q}\right)^2\right] \quad (17)$$

Left part: $$y_{aGauss} = a \cdot exp\left[-4ln2 \cdot \left(\frac{(x-p)}{2tq}\right)^2\right] \quad (18)$$

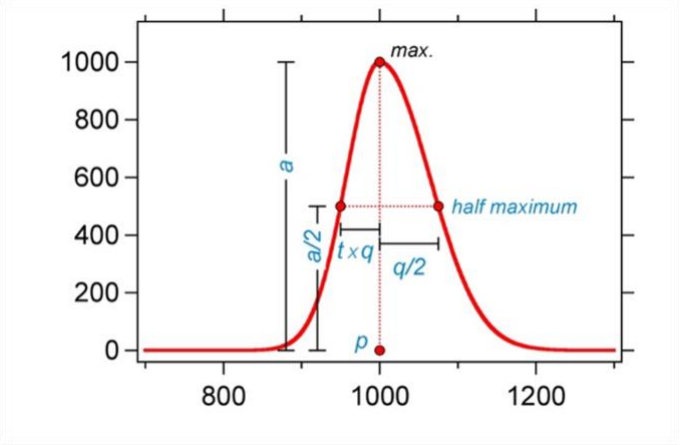

Figure S3. Example of an asymmetric Gaussian function with geometrical definitions of the parameters (*a* = 1000, *p* = 1000, *q* = 150, *t* = 0.3333, *FWHM* = 125).

S2.6 | Asymmetric Lorentzian function (LabSpec 6)

Right part: $$y_{aLorentz} = \frac{a}{1+4\left(\frac{x-p}{q}\right)^2} \quad (19)$$

Left part: $$y_{aLorentz} = \frac{a}{1+4\left(\frac{(x-p)}{2tq}\right)^2} \quad (20)$$

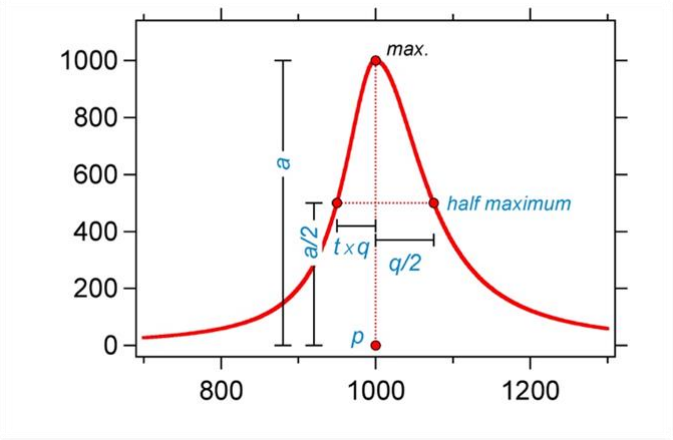

Figure S4. Example of an asymmetric Lorentzian function with geometrical definitions of the parameters (*a* = 1000, *p* = 1000, *q* = 150, *t* = 0.3333, *FWHM* = 125).

S2.7 | Asymmetric Logistic function (PeakFit v4.12)

$$y = a_0\left[1 + exp\left(-\frac{x + a_2 ln a_3 - a_1}{a_2}\right)\right]^{-a_3-1} {a_3}^{-a_3}(a_3+1)^{a_3+1} exp\left(-\frac{x + a_2 ln a_3 - a_1}{a_2}\right) \quad (21)$$

where a_0 is maximum amplitude, a_1 is centre position, a_2 is related to the width, and a_3 is the shape definition (>0).

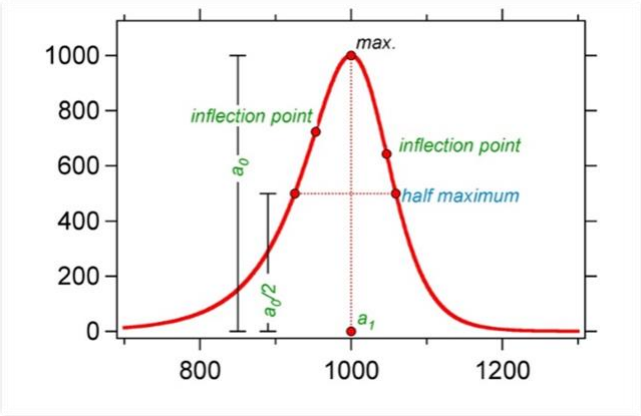

Figure S5. Example of an asymmetric Logistic function with geometrical definitions of the parameters (a_0 = 1000, a_1 = 1000, a_2 = 30, a_3 = 0.5, *FWHM* = 133.58).

The symmetric logistic function is obtained by the shape definition a_3 = 1

$$y = a_0\left[1 + exp\left(-\frac{x-a_1}{a_2}\right)\right]^{-2} 4\, exp\left(-\frac{x-a_1}{a_2}\right) \quad (22)$$

Alternative and equivalent definition:

$$f(x) = \frac{exp\left(-\frac{x-a}{b}\right)}{b\left[1+exp\left(-\frac{x-a}{b}\right)\right]^2} \quad (23)$$

inflexion points at: $x = a \pm ln\left(2+\sqrt{3}\right)\cdot b$ (24)

$$FWHM = 2\cdot ln\left(3+2\sqrt{2}\right)\cdot b \quad (25)$$

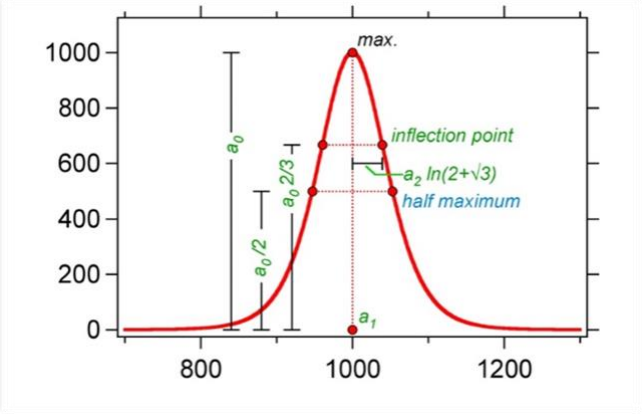

Figure S6. Example of a symmetric Logistic function with geometrical definitions of the parameters (a_0 = 1000, a_1 = 1000, a_2 = 30, *FWHM* = 105.76).

S3 | Spectrometer calibration of the LabRAM 300 and LabRAM HR Evolution

The transformation of a pixel number in the diode array of the detector to wavelength is calibrated with two parameters: 1. the position of the zero order of diffraction (0.0 nm); 2. the number of nm moved per motor step of rotation of the gratings by measuring a standard (either a known Raman shift or an atomic emission line). In other words, the general calibration method of the LabRAM system is a two-point fitting analysis with a linear interpolation which is reflected by these two parameters. The position of the zero order is defined by the parameter "zero" or "offset shift" in the software LabSpec 5 and 6, respectively, and it corresponds to the number of steps between the switch of the mechanical reference of the motor that drives the rotation and the position of 0 nm (i.e. zero order). Figure S7 illustrates that approximately 5 pixels are involved in the signal of the zero order in both systems. The LabRAM 300 system has a pixel resolution of 0.1393 nm at the zero order with the 600 mm^{-1} gratings (Fig. S7a) and the LabRAM HR Evolution 0.01105 nm with the 1800 mm^{-1} gratings (Fig. S7b). The pixel with the highest intensity is roughly optically estimated in a first approach, and subsequently positioned in the centre of a spectral window. It is recommended by Horiba to have this band position within ± 1 pixel of 0 nm (visual determination). The accuracy can be improved by estimation of the centre position with a symmetrical Gaussian-Lorentzian function. The latter procedure is performed by the automatic calibration option in LabSpec 6. Earlier versions of LabSpec offer the possibility to perform manually this fitting procedure. An example of the sensibility of the zero position to "offset shift" number is illustrated by a shift of 0.022 nm per 5 steps in the LabRAM 300 system with an 1800 mm^{-1} gratings, which is about half the size of one pixel.

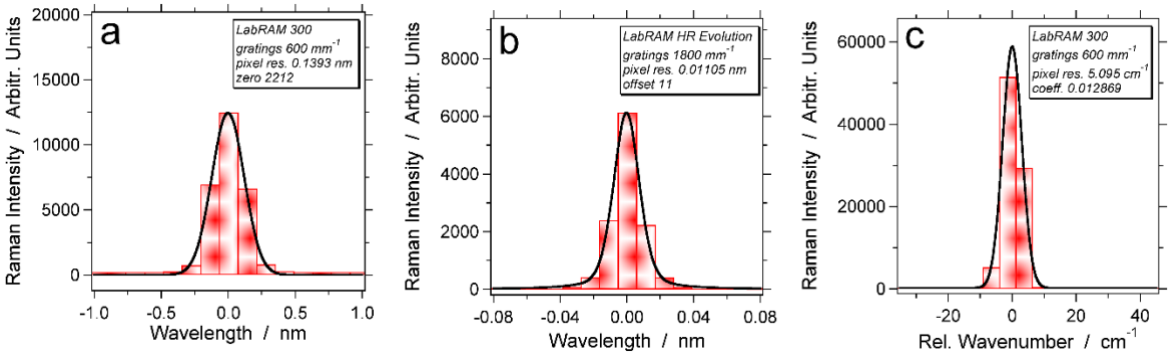

Figure S7. Spectra of the zero order (in wavelength, nm) of the LabRAM 300 system (**a**) and the LabRAM HR Evolution system (**b**). The pixel resolution (pixel res.) of the latter system is approximately improved by a factor ten. "zero" and "offset" numbers are instrument specific values. (**c**) Spectrum of the Nd-YAG laser (in relative wavenumber, cm^{-1}) obtained with the LabRAM 300 system. The coefficient (coeff.) is experimentally determined using a reference Raman band with a known relative wavenumber. The thick solid lines are symmetrical Gaussian-Lorentzian best-fit curves.

Subsequently, the Raman shift of silicon (Stokes scattering, $\Delta\nu$ = 520.7 ± 0.5 cm^{-1}), laser excitation wavelength (Rayleigh scattering, $\Delta\nu$ = 0.0 cm^{-1}; Fig. S7c), or an atomic emission lines (e.g. neon) is used to define the number of nm moved per motor step of the Sinus Arm Drive (i.e. "koeff" or "coeff" in LabSpec 5 and 6, respectively). The automatic calibration procedure allows the definition of only one specific Raman band as a relative wavenumber; therefore, neon lines must be recalculated in a virtual Raman shift number. For example, the 638.29914 nm neon line is transformed to 135.7462 cm^{-1} shift with the He-Ne laser (excitation laser wavelength 632.816 nm). The measured neon line or a reference Raman band is always positioned in the centre of a spectral window in the automatic calibration procedure. Similarly to the zero order, the peak position of the signal is determined with a symmetrical Gaussian-Lorentzian distribution function. Finally, a measurement session can start with the estimated "offset shift" and "coeff" values.

S4 | Uncertainty of gratings rotation in the calibration procedure

Automatic estimated peak positions of reference Raman bands in the calibration procedure of the LabSpec software are seldomly equal to the predefined values (Table S1). The LabSpec software allows a manual correction of "coeff" parameter to be able to reproduce accurate wavelength values. For example, the *540.05616 nm* neon line is initially estimated at 540.0510 nm, as defined by a Gaussian-Lorentzian best-fit peak and a coefficient of 0.002019020. This position (540.0562 nm, "First" in Table S1) is only obtained after manual fine-tuning of the coefficient to 0.002019029. This example illustrates that a relatively small difference in the coefficient (ca. 10^{-8}) may shift the initial detection of wavelength with 5 pm.

After the calibration procedure, rotation of the gratings forth and back, relocating the same spectral window position affects the reproducibility of the reference peak position ("Second" in Table S1). For example, a fine-tuned calibration of the *630.47893 nm* neon line with the LabRam 300 results in a peak position of 630.479 nm, however after gratings rotations this position is modified to 630.493 nm (i.e. "First" and "Second" in Table S1). This deviation corresponds to ca. 0.35 cm^{-1} wavenumber. The Sinus Arm Drive is not able to provide a sufficiently smooth movement to relocate the gratings with equal constant values. Peak positions of the same neon line calibrated with the newer LabRAM HR Evolution system may shift up to 3 pm, which is a significant improvement of the mechanical accuracy of the Sinus Arm Drive in more modern systems. Nevertheless, the Sinus Arm Drive adds an extra source of uncertainty in the estimation of calibrated peak position of Raman bands.

Table S1. Calibration of several neon lines (*calibration references*) according to the initial calibration procedure of the Raman spectrometers (*Automatic*), fine-tuned coefficients with highly accurate wavelengths (*First*), and the uncontrolled shift after rotation of the gratings with the Sinus Arm Drive (*Second*). See text for further details.

	Laser	*Gratings (mm^{-1})*	*Calibration reference (nm)*	*Automatic (nm)*	*Coefficient (fine-tuned)*	*First (nm)*	*Second (nm)*
LabRAM 300	Nd-YAG	1800	630.47893	630.487	0.00429046	630.479	630.493
LabRAM HR	He-Ne	1800	653.28824	653.262	0.00201911	653.2889	653.2858
Evolution	He-Ne	600	753.57739	753.589	0.006057263	753.5784	753.5668
	Nd-YAG	1800	540.05616	540.051	0.002019029	540.0562	540.0553
	He-Ne	1800	650.65277	650.642	0.002019090	650.6522	650.6564

S5 | Laser wavelength calibration

The well-defined He-Ne laser is selected to test the calibration procedures described in paragraph 6.3 to determine the wavelength of the laser (Fig. S8). The one-point calibration (see paragraph 6.1) is performed with the *638* neon line (Fig. S8a). This relatively distant neon line is located at about +5.5 nm higher wavelength than the laser, consequently it is expected to obtain slightly deviating values of the laser wavelength. The laser wavelength is estimated at values between 632.816 and 632.820 nm, with a mode at 632.816 nm according to multiple settings with variable positions of the spectral window. Subsequently, the laser wavelength is estimated with the proper bracketing technique using a best-fit linear correction through four adjacent neon lines: *626*, *630*, *633*, and *638* (Fig. S8b). The laser signal is approximately positioned in the centre of the spectral window. The calibrated wavelength of the He-Ne laser varies between 632.813 and 632.818 nm and is averaged at 632.816 ± 0.003 nm. Repetition of this procedure over several weeks resulted in similar values. In conclusion, the He-Ne laser wavelength in the LabRAM HR Evolution system is reproduced with the modified bracketing technique with a variability of 3 pm.

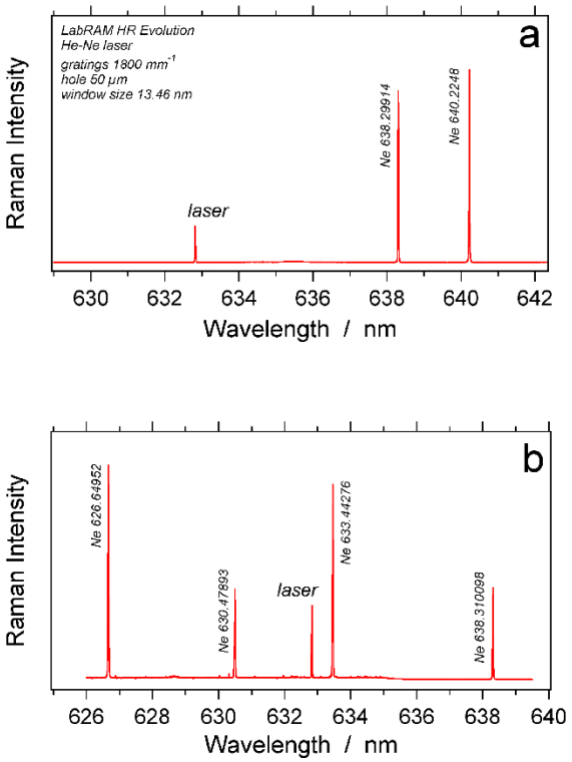

Figure S8. Spectrum of Rayleigh scattering of the He-Ne laser (632.816 nm) and two adjacent neon emission lines (*638* and *640*) in the Stokes scattering range (**a**), and with two adjacent neon lines (*626, 630, 633, and 6*38) on either side in both the Stokes and anti-Stokes scattering range (**b**) by positioning the neon lamp behind the edge filter. See text for further details.

An example of Nd-YAG laser wavelength determination is illustrated in Figure S9. The reported values given in the text and figures are only valid to this example due to the variability of the laser wavelength. The one-point calibration method (see paragraph 6.1) using the 534 neon line results in calibrated wavelengths between 532.044 to 532.048 nm (Fig. S9a). The *533* neon line that is closer positioned to the laser can also be used for the one-point calibration, resulting in a calibrated wavelength of 532.036 to 532.045 nm (Fig. S9b). This relatively weak neon line cannot be detected with the use of an edge filter and is mainly observable by positioning the neon lamp behind the edge filter. The relatively low intensity of the *532*, *533*, and *534* neon lines may prevent any reliable Gaussian-Lorentzian peak fitting results, and it may introduce relatively large uncertainties in peak positions. The use of neon lines closer to the Rayleigh scattering result in lower calibrated wavelengths, and the best approach results in a wavelength of ca. 532.035 nm. The variability of the laser wavelength is evidenced with similar

calibration procedures with the same instrument settings performed in the same laboratory after several days (532.055 to 532.058 nm), or weeks (532.090 to 532.102 nm, mode at 532.100 nm).

The proper bracketing technique (see paragraph 6.2) to calibrate the wavelength of the Nd-YAG laser involves six adjacent neon lines: *529*, *530*, *532*, *533*, *534*, and *534a* (Fig. S9b). The resulting calibrated wavelength varies significantly according to individual measurement sessions and is affected by variable room conditions. For example, a calibrated wavelength of 532.059 ± 0.005 nm could not be reproduced the following day, that revealed wavelengths of 532.113 ± 0.006 nm. Each session has a laser wavelength reproducibility of about 6 pm, i.e. approximately twice as much as the uncertainty of the He-Ne laser. The examples illustrate that the Nd-YAG laser must be experimentally determined before any measurement session, and that it is very unlikely that it has a constant wavelength over longer periods of investigation.

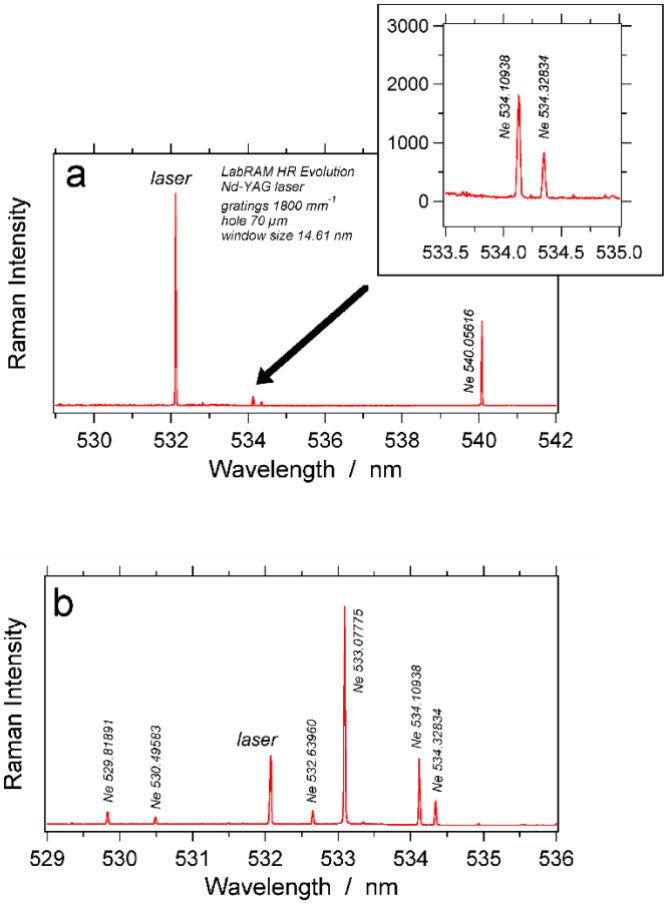

Figure S9. (**a**) Spectrum of the Rayleigh scattering of the Nd-YAG laser and three adjacent neon emission lines (*534*, *534a* and *540*) in the Stokes scattering range that can be used for the one-point calibration procedure. (**b**) the same spectrum after positioning the neon lamp behind the edge filter with six neon lines on both sides of the laser wavelength that can be used for calibration with the proper bracketing technique (Eq. 1 and 2).

S6 | Spectral window internal irregularities

The estimated peak positions of neon lines and Raman bands with Gaussian-Lorentzian best-fit probability distribution functions are significantly affected by their relative position in a spectral window (Fig. S10, S11).

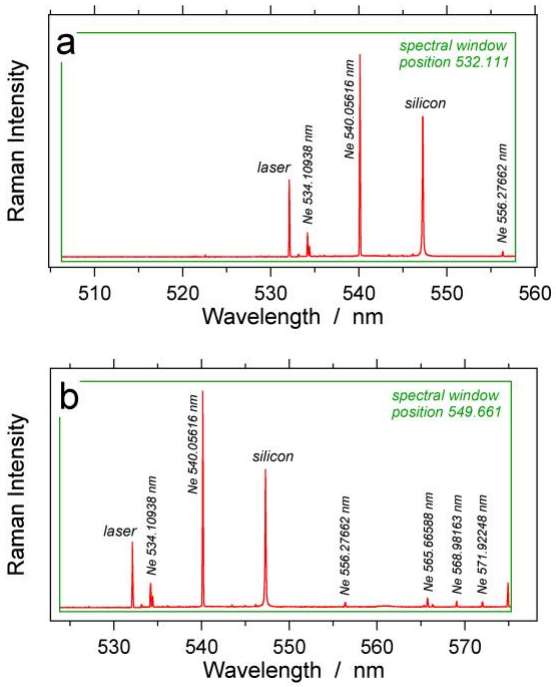

Figure S10. Raman band of silicon and neon emission lines in two spectra recorded in different spectral window position: **a**. 532.111 nm; and **b**. 549.661 nm. The silicon band is bracketed and calibrated by the 540.05616 nm and 556.27662 nm neon lines. This setup is used to determine the variability in peak position. Raman system: LabRAM HR Evolution, Nd-YAG laser, 600 mm^{-1} gratings, 70 mm confocal hole.

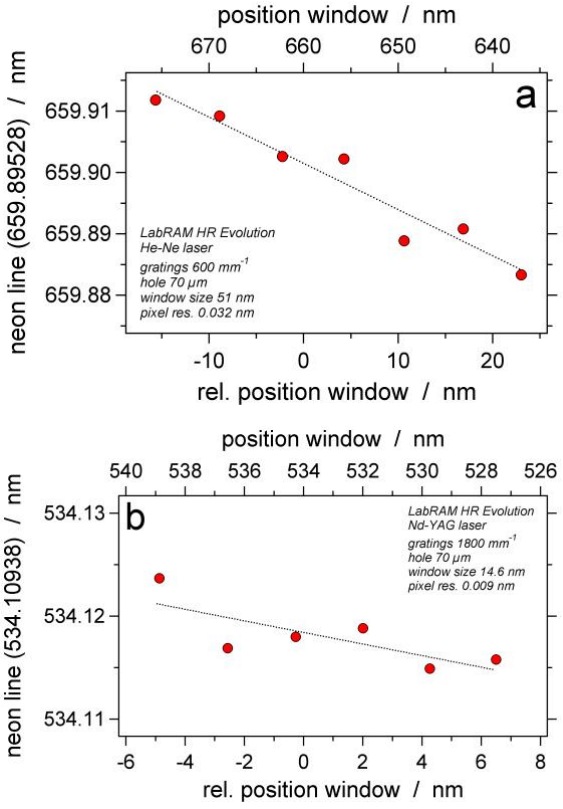

Figure S11. Shift in the detected peak positions of neon emission lines 659.89528 nm (**a**) and 534.10938 nm (**b**) dependent on their position within a single spectral window. The position is specified in absolute wavelength and in a relative number. The latter is defined by the relative position of the neon line within a spectral window (in nm): 0 is the centre position, positive values represent lines that appear to the right site, and negative values to the left site

S7 | Best-fit probability distribution functions

The variation in peak position estimation according to multiple *PDF's* decreases if the pixel resolution is increased by using a larger line density of the gratings (1800 mm^{-1} in Fig. S12a) and if more pixels are involved in the shape definition by using a larger focal length (Fig. S12b). The 574 neon line is detected by about 6 pixels with the LabRAM 300 (1800 mm^{-1} gratings) (Fig. S12a). All best-fit *PDF* have similar shapes, but the peak position may still vary significantly between 1398.220 cm^{-1} (Gaussian) to 1398.406 cm^{-1} (asymmetric Gaussian-Lorentzian). The Lorentzian curve result in the most deviating intensity values and must be regarded an unreliable shape reproduction of this neon line. The LabRAM HR Evolution with 1800 mm^{-1} gratings detects the 650 neon line within 12 pixels (Fig. S12b). The variation in peak estimation according to different *PDF* is decreased to about 0.03 cm^{-1}, i.e. from 433.192 cm^{-1} (Gaussian) to 433.208 cm^{-1} (asymmetric Gaussian-Lorentzian). This variation is similar to the variability that resulted from the precision of the laser wavelength (632.816 ± 0.001 nm), corresponding to a wavenumber variability of ± 0.025 cm^{-1}.

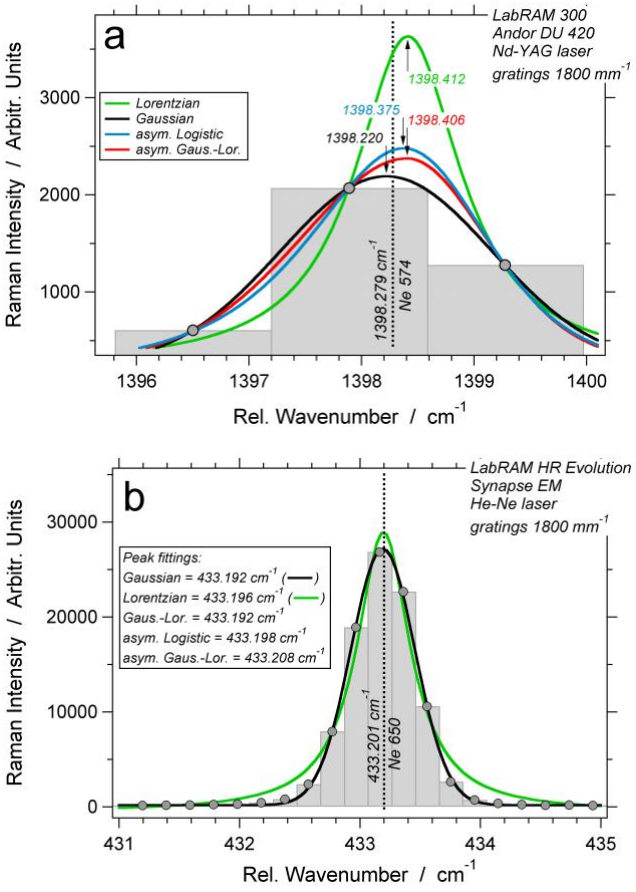

Figure S12. (**a**) Various *PDF* fitted to the 574 neon line analysed with the LabRAM 300 system (1800 mm^{-1} gratings). (**b**) Gaussian (black curve) and Lorentzian (green curve) distribution curves fitted to the 650 neon line with the LabRAM HR Evolution system. Peak positions (in cm^{-1}) of various *PDF's* are given in the included box.

S8 | Wandering neon emission line

The *633* neon line is analysed with a Nd-YAG laser (532.13 nm) in the LabRAM HR Evolution system and is shifted systematically to lower relative wavenumbers with increasing wavelength of the spectral window position (Fig. S13), i.e. the reversed effect compared to Figure 12b (with a He-Ne laser). The minimum step size of the gratings rotation is 0.002 nm ("coeff" parameter 0.002019074), and only 4 steps are included in one cycle due to a pixel resolution of 0.209 cm^{-1}. The symmetric and asymmetric Gaussian-Lorentzian best-fit curves reveal a gently wavy appearance, and corresponding peak positions are within the uncertainty of *SMT*, both at the upper and lower end of the error bars.

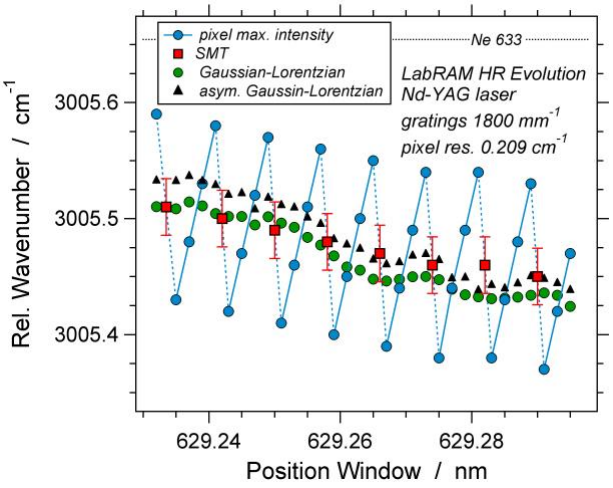

Figure S13. Relative wavenumber of the peak position of the 633 neon line as a function spectral window position.

S9 | Comparison with literature: Methane

The calibrated peak positions of CH_4 within the fluid inclusions illustrated in Figure 13 and the relationship to CH_4 density are inconsistent with any equation published in literature. To illustrate this controversy, the density of the 0.3461 g·cm^{-3} fluid inclusion (Fig. 13a) is calculated at 0.300 ± 0.007 g·cm^{-3} with the equations from Sublett et al. (2020) Moreover, the experimental data in literature reveal a large variability, and inconsistent relationship between density and Raman shift. The discrepancies were recognized by Lu et al. (2007), Lin et al. (2007), Shang et al. (2014), Zhang et al. (2016) and Sublett et al. (2020). The simple addition or subtraction of an arbitrary number ν_0 was suggested to equalize different data sets (Zhang et al., 2016; Lu et al., 2007), that reveal similar trends (slope) in the relationship between CH_4 Raman band and density (or pressure). Sublett et al. (2020) merely "corrected" their experimental data to obtain similar values as Lin et al. (2007). These studies do not adequately address the cause of the variation. Mathematical equations given by Lin et al. (2007), Shang et al. (2014), and Sublett et al. (2020) have no general application to fluid inclusion studies as they are instrument specific, i.e. they cannot be applied to Raman systems other than the ones that are used in these studies. The equation for the difference between CH_4 Raman bands "*near zero*" density and the measured shift (*D*) from Zhang et al. (2016) can also not be used in fluid inclusion research because *D* values between -3 cm^{-1} and -7.2 cm^{-1} can be assigned to multiple densities within the relatively large interval 0.11 to 0.53 g·cm^{-3}. For example, a *D* value of -5 cm^{-1} corresponds to a density of 0.1962 g·cm^{-3} and 0.4850 g·cm^{-3}. It must be noted that *D* equations always need an extra estimation, i.e. a hypothetical constant value of ν_0 to be able to calculate a density from a specific Raman shift.

The main causes of inconsistent data sets are highly underestimated uncertainty of individual measurements and partly insufficient calibration procedures, as described in the preceding paragraphs. Reproducibility, i.e. the average and standard deviation of multiple measurements of the same Raman band is often mistaken for the uncertainty in individual measurements (De Bièvre, 2008; Drosg, 2009). The CH_4 Raman bands measured with different Raman systems in the present study reveal consistent peak positions and it is expected that similar calibration procedures applied to previously work would have resulted in a consistent data set, and a uniform equation of the relationship between Raman band peak position of CH_4 and density, however, with a much larger error as indicated in previous publications.

References

De Bièvre P (2008) Measurement uncertainty is not synonym of measurement repeatability or measurement reproducibility. Accred. Qual. Assur., vol. 13, 61-62

Drosg M (2009) *Dealing with uncertainties, a guide to error analysis*, Springer, Heidelberg.

Lin F, Bodnar RJ, Becker SP (2007) Experimental determination of the Raman CH_4 symmetric stretching (ν_1) band position from 1–650 bar and 0.3–22 °C: Application to fluid inclusion studies. Geochimica et Cosmochimica Acta, vol. 71, 3746-3756.

Lu W, Chou I-M, Burruss RC, Song Y (2007) A unified equation for calculating methane vapor pressures in the CH_4–H_2O system with measured Raman shifts. Geochimica et Cosmochimica Acta 71, 3969–3978.

Shang L, Chou I-M, Burruss RC, Hu R, Bi, X (2014) Raman spectroscopic characterization of CH_4 density over a wide range of temperature and pressure. Journal of Raman Spectroscopy, vol. 45, 696-702.

Sublett DM jr., Sendula E, Lamadrid H, Steele-MacInnes M, Spiekermann G, Burruss RC, Bodnar RJ (2020) Shift in the Raman symmetric stretching band of N_2, CO_2, and CH_4 as a function of temperature, pressure, and density. Journal of Raman Spectroscopy, vol. 51, 555-568.

Zhang J, Qiao S, Lu W, Hu Q, Chen S, Liu Y, 2016. An equation for determining methane densities in fluid inclusions with Raman shifts. J. Geochem. Explor. 171, 20–28.

S10 | CO_2 density estimation of fluid inclusions

At room temperature, the fluid inclusions in Figure 16 contain an aqueous phase and a homogeneous CO_2-rich phase. Microthermometry reveal melting temperatures of CO_2 at -56.6 ˚C in the presence of ice and CO_2-clathrate. The CO_2 phases in the fluid inclusion in Fig. 16a homogenize at +12.6 ˚C in the vapour phase. Despite optical difficulties to estimate this temperature, cycling experiments with the use of metastability due to nucleation difficulties in cooling systems result in an uncertainty of ±0.1 ˚C (i.e. best value of a heating-freezing stage). CO_2-clathrate dissolution occurs at + 9.9 ˚C (Q2 melting) (Bakker, 1997). The density of the homogeneous CO_2 phase can be directly calculated with the equation of state from Span and Wagner (1996) that was adapted for fluid inclusion work in the programs *Loner18* and *LonerSpW* in the software package *FLUIDS* (https://fluids.unileoben.ac.at) (Bakker, 2003) for pure CO_2 phases at 0.1477 ± 0.0006 g·cm^{-3} (298 ± 1 cm^3·mol^{-1}). The low concentration of H_2O in the vapour bubble, i.e. 0.001 mole% at room temperatures is assumed to have a negligible effect on Raman peak properties of CO_2. Fluid inclusion in Fig. 16b reveals a homogenization of CO_2-rich phases at +9.3 ˚C in the liquid phase in the presence of a CO_2-clathrate and an aqueous liquid solution. The clathrate dissolves completely at +10.1 ˚C. The CO_2 liquid and vapour phases homogenize at +6.2 ˚C to a liquid phase in the metastable absence of CO_2-clathrate and in the presence of only an aqueous liquid solution. The dissolution and homogenization behaviour of this fluid inclusion in a heating experiment corresponds to settings described in Diamond (1992) and Bakker (1997), and its fluid properties can be modelled with the program "*NOSALT*" and "*CURVES*" in the package "*Clathrates*" (Bakker, 1997). The density of the homogeneous CO_2 phase can be directly modelled with the homogenization temperature in the metastable absence of a clathrate using the same programs *Loner18* and *LonerSpW*: 0.8880 ± 0.0007 g·cm-3 (49.56 ± 0.03 cm^3·mol^{-1}). This vapour phase contains 0.4 mole% H_2O at the clathrate dissolution temperature and is considered to behave as a pure CO_2 phase.

References

Bakker RJ .(1997) CLATHRATES: computer programs to calculate fluid inclusions V-X properties using clathrate melting temperatures. Computers and Geosciences vol. 23, 1-18.

Bakker RJ (2003) Package FLUIDS 1. Computer programs for analysis of fluid inclusion data and for modelling bulk fluid properties. Chemical Geology, vol.194, 3-23.

Diamond LW (1992) Stability of CO_2 clathrate hydrate + CO_2 liquid + CO_2 vapour + aqueous KCI-NaCl solutions: Experimental determination and application to salinity estimates of fluid inclusions: Geochim. Cosmochim. Acta, v. 56, p. 273-280.

Span R, Wagner W (1996) A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. J. Phys. Chem. Ref. Data 25, 1509–1596.

S11 | Comparison with literature: Carbon Dioxide

Calibrated CO_2 Raman bands and corresponding densities are compared to the purely empirical equations presented in previous publications (Table S2). A large variety of the Fermi diad distance is calculated for a measured density of 0.1477 g·cm^{-3} (fluid inclusion Fig. 16a): from 102.828 to 103.202 cm^{-1}. Vice versa, the density that is calculated with numerous equations presented in literature using a calibrated Fermi diad of 103.123 cm^{-1} is varying between 0.1238 and 0.2623 g·cm^{-3}, corresponding to a relative difference up to 78% (Table S2a). The calculated densities for the 0.8880 g·cm^{-3} fluid inclusion (Fig. 16b) vary to an even larger extend, from 0.7517 to 0.9463 g·cm^{-3} (Table S2b). The main causes of these differences are already elucidated for the CH_4 examples, and include inadequate calibration methods and underestimated uncertainties of individual measurements. Raman systems that were calibrated with reference Raman bands of naphthalene, benzonitrile, and diamond do not provide simultaneously recorded CO_2 spectra (Kawakami et al., 2003; Yamamoto and Kagi, 2006; Wang et al., 2011; Remigi et al., 2021). These studies ignore the uncertainty of the reference values: for example ±0.29 to ±0.49 cm^{-1} for naphthalene, and ±0.39 to ±0.98 cm^{-1} for benzonitrile (McCreery, 2000). The inferred uncertainty of the Fermi diad based on these reference bands does not allow the estimation of an accurate equation to define peak positions within the wavelength range detected by one pixel. These uncertainties often exceed pixel resolutions, for example ca. 1.5 cm^{-1} in Kawakami et al. (2003), Yamamoto and Kagi (2006); ca. 1 cm^{-1} in Wang et al. (2011).

Table S2a. Comparison of density and Fermi diad distances of the fluid inclusion illustrated in Figure 16a with published purely empirical equations. Pressure in fluid inclusion at 22 °C is 5.21 MPa.

Reference	*Fermi diad at 0.1477 g/cm^3*	*density at 103.123 cm^{-1}*	*density difference*
	cm^{-1}	*g/cm^3*	*g/cm^3*
Rosso and Bodnar (1995)	103.048	0.1779	+0.0302 (20%)
Kawakami et al. (2003)	102.828	0.2623	+0.1146 (78%)
Yamamoto and Kagi (2006)	102.868	0.2339	+0.0862 (58%)
Song et al. (2009)	103.033	0.1864	+0.0387 (26%)
Wang et al. (2011)	103.103	0.1559	+0.0082 (6%)
Fall et al. (2011)	103.035	0.1843	+0.0366 (25%)
Lamadrid et al. (2017)	103.008	0.1887	+0.0410 (28%)
Wang et al. (2019) eq. 3	103.182	0.1255	-0.0222 (15%)
Wang et al. (2019) eq. 2	103.113	0.1507	+0.0030 (2%)
Le et al. (2019), a	103.186	0.1238	-0.0239 (16%)
Le et al. (2019), c	103.202	0.1242	-0.0235 (16%)
Sublett et al. (2019)	103.043	0.1836	+0.0359 (24%)
Remigi et al. (2021)	103.060	0.1849	+0.0372 (25%)

Line segment normalization with simultaneously recorded neon lines (cf. Fig. 2b) was performed by Lamadrid et al. (2017), Wang et al. (2019)[a]. Line segment calibrations are only able to correct the difference between two Raman bands, e.g. the Fermi diad. The upper and lower CO_2 band was bracketed with two neon lines by Lamadrid et al. (2017), whereas Wang et al. (2019) use two neon lines both at higher relative wavenumbers than the upper CO_2 band, i.e. outside the wavenumber range where the Fermi diad occurs. The two methods resulted in significantly different density estimations

[a] Footnote: Wang et al. (2019) contains many typographical errors in the definition of parameters: the minus signs are missing for the parameters a14, a16, a19, a22, a26, and a29 in their equations 2 and 3.

(Table S2). Moreover, Wang et al. (2019) provide two inconsistent equations to calculated density from the Fermi diad distance (e.g. 103.123 cm^{-1}, Table S2a): 0.1255 $g \cdot cm^{-3}$ using temperature (Eq. 3 in Wang et al., 2019), and 0.1507 $g \cdot cm^{-3}$ using pressure (Eq. 2 in Wang et al., 2019). Both equations can be used assuming that the fluid inclusion in Figure 16a contains at 22 ˚C a CO_2-rich vapour bubble with an internal pressure of 5.21 MPa (calculated with Span and Wagner, 1996). The calibration method of a specific Raman band according to Lin et al. (2007) was used by Fall et al. (2011) and Sublett et al. (2020), using two neon lines that bracket the Fermi diad, and both studies reveal similar results. However, both the upper and lower band of CO_2 are corrected with equal values according to this method (cf. paragraph 4 “Calibration method bazar”), i.e. a correction value defined at the centre between the two neon lines (1031.42 and 1458.58 cm^{-1}). Consequently, the Fermi diad is not affected by these corrections, which correspond to the assumption of a “linear offset” of the monochrometer. Unfortunately, Fall et al. (2011)] and Sublett et al. (2020) do not provide equations for the relationship between individual upper or lower CO_2 band peak positions and density or pressure.

Table S2b. Comparison of density and Fermi diad distances of the fluid inclusion illustrated in Figure 16b with published purely empirical equations. Pressure in fluid inclusion at 22 °C is 14.38 MPa.

Reference	*Fermi diad of 0.8880 g/cm³*	*density of 104.709 cm⁻¹*	*density difference*
	cm^{-1}	g/cm^3	g/cm^3
Rosso and Bodnar (1995)	104.891	0.8149	-0.0731 (8%)
Kawakami et al. (2003)	104.558	0.9463	+0.0583 (7%)
Yamamoto and Kagi (2006)	104.604	0.9240	+0.0360 (4%)
Song et al. (2009)	104.772	0.8639	-0.0241 (3%)
Wang et al. (2011)	104.773	0.8654	-0.0226 (3%)
Fall et al. (2011)	104.773	0.8649	-0.0231 (3%)
Lamadrid et al. (2017)	105.093	0.7517	-0.1363 (15%)
Wang et al. (2019) eq. 3	104.844	0.8349	-0.0531 (6%)
Wang et al. (2019) eq. 2	104.853	0.8343	-0.0537 (8%)
Le et al. (2019), a	104.846	0.8362	-0.0518 (6%)
Sublett et al. (2019)	104.836	0.8396	-0.0484 (5%)
Remigi et al. (2021)	104.541	0.9484	+0.0604 (7%)

References

Fall A, Tattitch B, Bodnar RJ (2011). Combined microthermometric and Raman spectroscopic technique to determine the salinity of H_2O-CO_2-NaCl fluid inclusions based on clathrate melting. Geochim. Cosmochim. Acta 75, 951–964.

Kawakami Y, Yamamoto J, Kagi H (2003) Micro-Raman Densimeter for CO_2 Inclusions in Mantle-Derived Minerals. Applied Spectroscopy, vol. 57, 1333-1339.

Lamadrid HM, Moore lR, Moncada D, Rimstidt JD, Burruss RC, Bodnar RJ (2017) Reassessment of the Raman CO_2 densimeter. Chemical Geology, vol. 450, 210-222.

Le V-H, Caumon M-C, Tarantola A, Randi A, Robert P, Mullis J (2019) Quantitative Measurements of Composition, Pressure, and Density of Microvolumes of CO_2–N_2 Gas Mixtures by Raman Spectroscopy. Analytical Chemistry, vol. 91, 14359-14367.

Lin F, Bodnar RJ, Becker SP (2007) Experimental determination of the Raman CH_4 symmetric stretching (v_1) band position from 1–650 bar and 0.3–22 °C: Application to fluid inclusion studies. Geochimica et Cosmochimica Acta, vol. 71, 3746-3756.

McCreery RL (2000) Raman Spectroscopy for Chemical Analysis. John Wiley & Sons.

Remigi S, Mancini T, Ferrando S, Frezzotti M-L (2021) Interlaboratory application of Raman CO_2 densimeter equations: experimental procedure and statistical analysis using bootstrapped confidence intervals. Applied Spectroscopy, vol. 75, 867-881.

Rosso KM, Bodnar RJ (1995) Microthermometric and Raman spectroscopic detection limits of CO_2 in fluid inclusions and the Raman spectroscopic characterization of CO_2. Geochimica et Cosmochimica Acta, vol. 59, 3961-3975.

Song Y, Chou, I-M, Hu W, Burruss R, Lu W (2009) CO_2 density-Raman shift relation derived from synthetic inclusions in fused silica capillaries and its application. Acta Geologica Sinica, vol. 83, 932-938.

Span R, Wagner W (1996) A new equation of state for carbon dioxide covering the fluid region from the triple-point temperature to 1100 K at pressures up to 800 MPa. J. Phys. Chem. Ref. Data 25, 1509–1596.

Sublett DM jr., Sendula E, Lamadrid H, Steele-MacInnes M, Spiekermann G, Burruss RC, Bodnar RJ (2020) Shift in the Raman symmetric stretching band of N_2, CO_2, and CH_4 as a function of temperature, pressure, and density. Journal of Raman Spectroscopy, vol. 51, 555-568.

Wang X, Chou I-M, Hu W, Burruss RC, Sun Q, Song Y (2011) Raman spectroscopic measurements of CO_2 density: Experimental calibration with high-pressure optical cell (HPOC) and fused silica capillary capsule (FSCC) with application to fluid inclusion observations. Geochim. Cosmochim. Acta 75, 4080–4093.

Wang W, Caumon M-C, Tarantola A, Pironon J, Lu W, Huang Y (2019) Raman spectroscopic densimeter for pure CO_2 and CO_2-H_2O-NaCl fluid systems over a wide P-T range up to 360 °C and 50 MPa. Chemical Geology, vol. 528.

Yamamoto J, Kagi H (2006) Extended Micro-Raman Densimeter for CO_2 Applicable to Mantle-originated Fluid Inclusions. Chemistry Letters, vol. 35, 610-611.