Case study I: Modelling refractive index dispersion

Adam Stone
Department of Materials Science and Engineering, Lehigh University, Bethlehem, PA 18015, USA
(Dated: June 27, 2011)

The Levenberg-Marquardt algorithm is used to fit refractive index dispersion data of an unknown sample. A suitable model based on the Sellmeier equation is chosen by examining χ^2 values, residuals, and parameter correlations. The temperature dependence of the fitting parameters is then discussed.

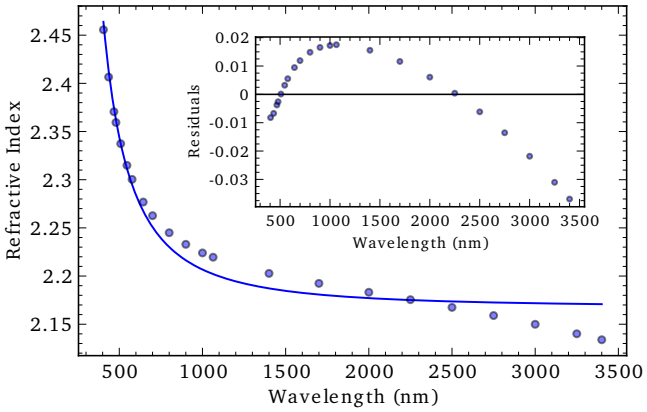

FIG. 1. Fit and residuals for one oscillator.

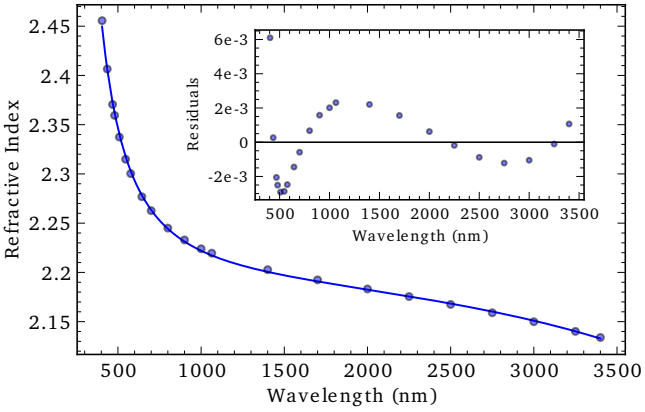

FIG. 2. Fit and residuals for two oscillators.

Refractive index data collected at several temperatures as a function of wavelength was provided for an unknown sample. This data was analyzed by curve-fitting with a Levenberg-Marquardt algorithm and model equation based on the Sellmeier equation (Eq. 1).

$$n = \sqrt{1 + \sum_{i=1}^{m} \frac{S_i}{1 - (\lambda_i/\lambda)^2}} \qquad (1)$$

This equation relates refractive index, n, to a series of m oscillators of strength S_i centered at wavelength λ_i. In the simplest case, we can try to fit the data with a single oscillator. The result is shown in Fig. 1, with an oscillator at $\lambda_1 = 211 \pm 3$ nm (where $\pm$ indicates the standard deviation). The fit looks poor, with $\chi^2 = 0.00494$, and the residuals are systematic rather than random scatter. A single oscillator model appears insufficient.

Adding a second oscillator substantially improves the fit, yielding oscillators at $\lambda_1 = 199 \pm 1$ and $\lambda_2 = 6450 \pm 1070$ nm. The χ^2 drops to 9.77×10^{-5}, and the curve that qualitatively looks much closer to the raw data (Fig 2). However, the error in λ_2 is undesirably high, and the residuals still show systematic error which suggests that the model can be further improved.

We can continue along the same path by adding a third oscillator. This fit has a χ^2 of 6.61×10^{-8}, which seems like a substantial improvement, and the residuals look more like random error than a deficiency in the model. The oscillators are now at $\lambda_1 = 142 \pm 5$, $\lambda_2 = 25400 \pm 3300$, and $\lambda_3 = 265 \pm 3$ nm, however, and both the position of λ_2 and its large error are cause for concern. An oscillator at such long wavelengths is not physically meaningful.

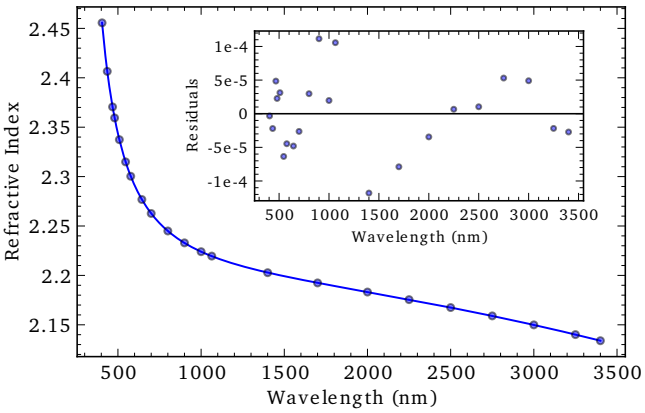

FIG. 3. Fit and residuals for three oscillators.

Covariance and Correlation

If the linear correlation between the strength and position of this IR oscillator is large, a stronger oscillator at very long wavelengths would have nearly the same

effect on the fit as a weaker oscillator at a more reasonable wavelength. We can quantify this effect using the covariance matrix returned by the Levenberg-Marquardt algorithm.

The diagonal components of the covariance matrix represent the variance of the parameters themselves, the square root of which gives their standard deviations. The off-diagonal elements indicate the covariance, which is related to the correlation between different parameters. The units of covariance are the product of the units of the individual parameters, and the unitless correlation is obtained by dividing the covariance by the standard deviation of each of the parameters (obtained from the diagonal elements). This allows more direct comparison between different parameter correlations by accounting for differences in units. Using this relation, the following correlation matrix is obtained:

$$\begin{pmatrix} & S_1 & \lambda_1 & S_2 & \lambda_2 & S_3 & \lambda_3 \\ S_1 & 1 & 0.998 & -0.598 & -0.596 & -0.999 & 0.998 \\ \lambda_1 & & 1 & -0.626 & -0.624 & -0.998 & 0.994 \\ S_2 & & & 1 & 0.999 & 0.600 & -0.580 \\ \lambda_2 & & & & 1 & 0.598 & -0.578 \\ S_3 & & & & & 1 & -0.998 \\ \lambda_3 & & & & & & 1 \end{pmatrix}$$

Because the correlation matrix is symmetric, it can be displayed with only one set of off-diagonal elements to avoid redundancy. Values closer to 1 in magnitude indicate stronger linear correlation, while the sign indicates the direction of the correlation. Focusing on these off-diagonal elements, several strong correlations are seen; not only between λ_2 and S_2, which was suspected, but between many of the other parameters as well. This is likely due to the fact that the correlation matrix doesn't take into account cross-correlations between other parameters that may influence the correlation of interest. For example, changes in one oscillator may compensate for changes in another, affecting the apparent correlation between the individual S_i and λ_i parameters of both.

A better basis for comparison may be the partial correlations, which give the linear correlation between two parameters *after excluding the influence of the other parameters*. To calculate the partial correlation between two parameters x and y while controlling for parameters z_1 ... z_n, involves a recursive formula:

$$r_{xy.z_1\ldots z_n} = \frac{r_{xy.z_1\ldots z_{n-1}} - r_{xz_n.z_1\ldots z_{n-1}} r_{yz_n.z_1\ldots z_{n-1}}}{\sqrt{(1 - r^2_{xz_n.z_1\ldots z_{n-1}})(1 - r^2_{yz_n.z_1\ldots z_{n-1}})}} \quad (2)$$

Examining this expression, each of the r terms applies the same function but with one fewer control parameters. Each of these r terms thus depends on three more r terms, but again with one fewer control parameters, and this recursion continues until reaching r terms with only one control parameter. At that point, these terms can be evaluated according to:

$$r_{xy.z_1} = \frac{r_{xy} - r_{xz_1} r_{yz_1}}{\sqrt{(1 - r^2_{xz_1})(1 - r^2_{yz_1})}} \quad (3)$$

Where each r_{xy} is the correlation between parameters x and y obtained from the correlation matrix. Eqs. 2 and 3 describe a single partial correlation between two parameters, but applying it to all parameters used in the fit yields a partial correlation matrix. For the three oscillator fit, the partial correlation matrix is:

$$\begin{pmatrix} & S_1 & \lambda_1 & S_2 & \lambda_2 & S_3 & \lambda_3 \\ S_1 & 1 & -0.696 & 0.629 & -0.627 & -0.991 & -0.618 \\ \lambda_1 & & 1 & 0.071 & -0.070 & -0.784 & -0.981 \\ S_2 & & & 1 & 0.999 & 0.561 & 0.049 \\ \lambda_2 & & & & 1 & -0.559 & -0.048 \\ S_3 & & & & & 1 & -0.716 \\ \lambda_3 & & & & & & 1 \end{pmatrix}$$

The strong (0.999) correlation remains between λ_2 and S_2 even after accounting for the other parameters, so the initial suspicion was correct. Most of the other strong correlations have been substantially reduced after controlling for the other parameters. The exceptions are the partial correlations between S_1 and S_3, and between λ_1 and λ_3, which are still large but inversely correlated. Since these are both close to each other in the UV range, it is reasonable that shifting one in position or strength could be compensated by an opposing shift in the other.

Large correlations between some parameters may be inevitable features of the model, and we have seen that two oscillators in the UV produces a substantially better fit than one while remaining physically sensible, even if the correlations between the two are strong. Large correlations can also suggest redundancy, however, if the model contains parameters which are effectively not contributing to the fit. Examining the Sellmeier equation, there is good reason to suspect this is the case for the IR oscillator at unreasonably long wavelengths. As λ becomes large, $(\lambda_i/\lambda)^2$ becomes much larger than 1 and the expression for the oscillator approximates to Eq. 4:

$$\frac{S_i}{1 - (\lambda_i/\lambda)^2} \approx -\frac{S_i}{\lambda_i^2}\lambda^2 \quad (4)$$

In this approximation S_i and λ_i are fully coupled, and changes in one can be completely offset by changes in the other. They effectively act as a single parameter $D_i = S_i/\lambda_i^2$. Although some information about the IR oscillator seems to be lost by doing so, we can substitute Eq. 4 into the fitting model for one of the three oscillators in order to gain the improved fit of a three oscillator model while reducing the correlation. The loss of specific

TABLE I. Fitting results for modified 3 oscillator model.

S_1	λ_1(nm)	S_2	λ_2(nm)	D(nm^{-2})
1.34 ± 0.12	258 ± 3	2.50 ± 0.12	129 ± 7	$(2.52 \pm 0) \times 10^{-8}$

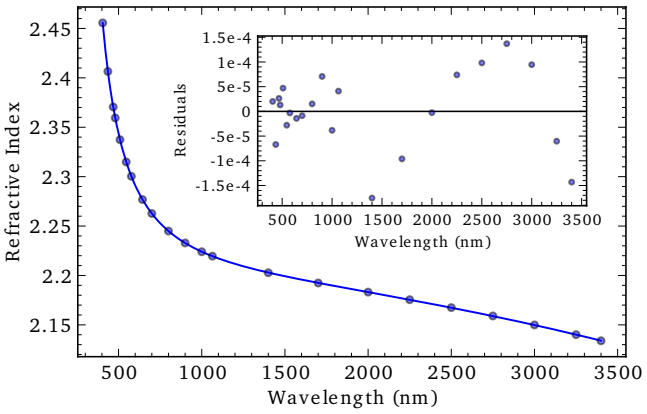

FIG. 4. Fit and residuals for three oscillators with Eq. 4 approximation used for the IR oscillator.

values for λ_2 and S_2 by combining them into one parameter is not really meaningful, because the high correlation was causing spurious results anyway.

The resulting fit after making this approximation is shown in Fig. 4. The two UV oscillators are located at $\lambda_1 = 258$ and $\lambda_2 = 129$ nm, while the location of the IR oscillator is now uncertain. The χ^2 is 1.24×10^{-7}, slightly larger than without this approximation, but the quality of fit is not necessarily better based on the value of χ^2 alone. In this case it is more sensible to describe the IR oscillator with one parameter. The final values obtained from this fit are given in Table I, and the partial correlation matrix of this model is given by:

$$\begin{pmatrix} & S_1 & \lambda_1 & S_2 & \lambda_2 & D \\ S_1 & 1 & -0.714 & -0.991 & -0.781 & 0.568 \\ \lambda_1 & & 1 & -0.619 & -0.982 & 0.051 \\ S_2 & & & 1 & -0.695 & 0.634 \\ \lambda_2 & & & & 1 & 0.073 \\ D & & & & & 1 \end{pmatrix}$$

Temperature Dependence

With a fitting model established, the data sets at different temperatures can be fit. If the model is representing the underlying physics, we would expect the parameters to show systematic changes as the temperature is increased. Performing these fits and plotting the parameter values vs. temperature yields Figs. 5, 6, and 7.

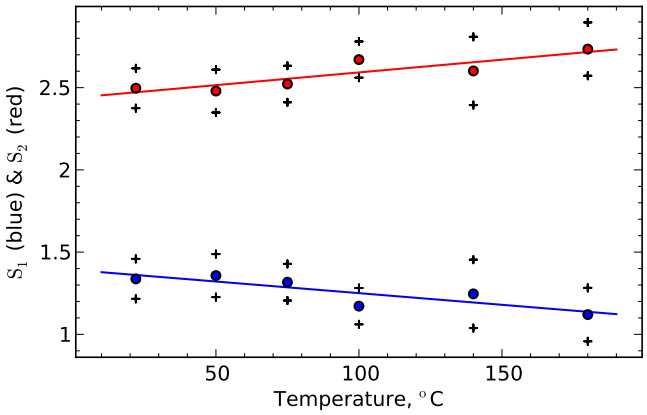

FIG. 5. Temperature dependence of S_1 and S_2.

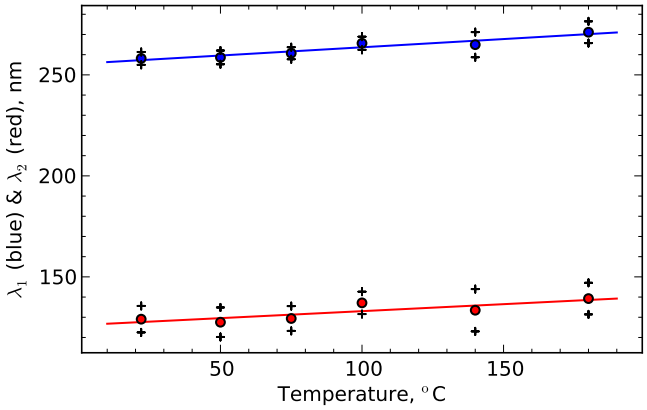

FIG. 6. Temperature dependence of λ_1 and λ_2.

The S and λ terms show approximately linear changes with temperature, with the error range indicating the parameter standard deviation obtained from the covariance matrix. While both oscillator positions increase comparably, the strengths trend in opposite directions. For the

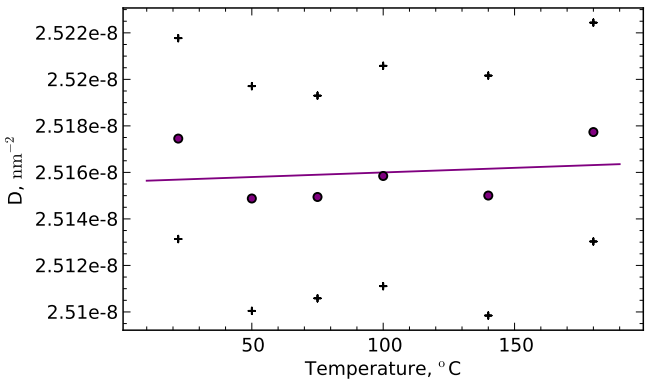

FIG. 7. Temperature dependence of D.

third oscillator, it is less clear whether a real trend occurs. Since this parameter is an approximation of S_3/λ_3^2, trends in either parameter could be lost in the ratio.

Conclusions

Fitting refractive index dispersion data is a challenge, because there are many variations in the Sellmeier model which can subtly influence the fit. Because the positions of the oscillators lie outside the data range, the most characteristic feature of the oscillator function (the discrete step at the oscillator position) are not represented in the fit. This can lead to misleading results if the strength and position of oscillators are strongly correlated, and means in practice that some information about oscillator strength and position may simply be inaccessible by this method. Likewise, the large correlations seen between oscillators in the UV range cast some doubt on their precision as well. However, the temperature trends in the UV parameters do lend some confidence to these results.