Case study I: Modelling a refractive index dispersion

Ivan Biaggio
Department of Physics, Lehigh University, Bethlehem, PA 18015, USA
(Dated: May 31, 2011)

This is a discussion of a first "case study": we have a data set describing the refractive index of a material at various temperatures, and we want to see how we can model it to both describe the empirical data well, and to extract physical parameters. Several activities are suggested, and this example will be useful to implement a few of the techniques that we will discuss later on.

Table I shows the data obtained for the refractive index of some material at different temperatures. There are two aims for this case study:

1. Find out if we can represent the refractive index data with a simple mathematical formula.
2. See if it is possible to extract some interesting physical parameter.

The first thing to do is to come up with an appropriate description for a refractive index as a function of the wavelength. One of the most used descriptions is based on the *Sellmeier Formula* [1]. In its simplest form it can be written as

$$n(\lambda)^2 - 1 = \frac{S_1}{1-(\lambda_1/\lambda)^2}, \quad (1)$$

TABLE I. Refractive index data at various temperatures. The maximum experimental error for a single data point is 3×10^{-4}

λ [nm]	22°C	50°C	75°C	100°C	140°C	180°C
404.66	2.4556	2.4584	2.4611	2.4641		
435.83	2.4065	2.4087	2.4108	2.4131	2.4173	2.4224
467.82	2.3705	2.3723	2.3740	2.3759	2.3793	2.3835
479.99	2.3594	2.3611	2.3627	2.3645	2.3677	2.3716
508.58	2.3374	2.3389	2.3403	2.3420	2.3449	2.3484
546.07	2.3149	2.3163	2.3175	2.3190	2.3215	2.3247
576.96	2.3003	2.3016	2.3027	2.3041	2.3065	2.3094
643.85	2.2768	2.2779	2.2789	2.2801	2.2822	2.2847
700.00	2.2627	2.2636	2.2645	2.2657	2.2676	2.2699
800.00	2.2450	2.2458	2.2466	2.2476	2.2494	2.2515
900.00	2.2329	2.2337	2.2344	2.2354	2.2370	2.2390
1000	2.2240	2.2248	2.2255	2.2266	2.2281	2.2299
1064	2.2195	2.2202	2.2209	2.2219	2.2232	2.2252
1400	2.2027	2.2035	2.2042	2.2049	2.2064	2.2081
1700	2.1924	2.1932	2.1937	2.1945	2.1959	2.1976
2000	2.1832	2.1839	2.1845	2.1852	2.1866	2.1882
2250	2.1755	2.1762	2.1768	2.1775	2.1789	2.1805
2500	2.1675	2.1682	2.1688	2.1695	2.1709	2.1725
2750	2.1590	2.1598	2.1603	2.1611	2.1624	2.1640
3000	2.1499	2.1507	2.1512	2.1519	2.1533	2.1549
3250	2.1401	2.1408	2.1415	2.1422	2.1435	2.1451
3400	2.1339	2.1346	2.1352	2.1359	2.1373	2.1389

where S_1 and λ_1 represent the strength and the position of an "oscillator", which often simply stands in for an electronic transition of certain strength (the oscillator strength S_1) that happens for a given photon energy $\hbar c/\lambda_1$ in the high energy region beyond the transparency region of the material.

Figure 1 shows a plot of the refractive index data, with the inset showing the not extremely big, but significant change in the data as the temperature is raised.

Simplest approach

In order to investigate how this data can be modeled, I suggest that you first *fit* the room temperature refractive index data in the first column of Table I using Eq. (1. Such a fit will deliver two numbers, the values of the parameters S_1 and λ_1. What values do you get? In which spectral region is λ_1? Do you think this is a good fit? Discuss how good Eq. (1) is for fitting the refractive index data. What are the limitations?

Adding parameters to improve a fit

Next, check what happens if you extend Eq. (1) so that it doesn't consist of only one oscillator, but two! This is a natural extension. Fit again. Look at the result. Is it better now? Is it good enough? Is there something else

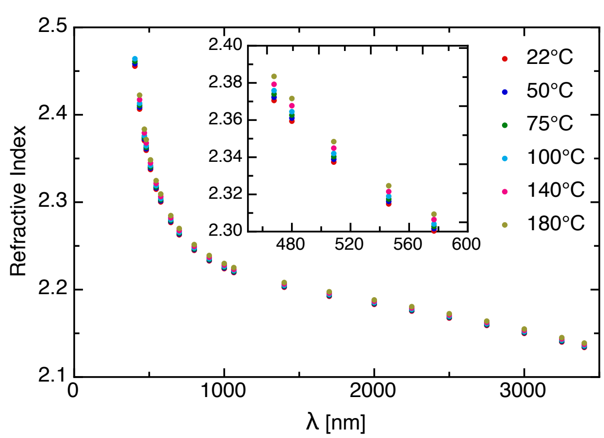

FIG. 1. Graphical representation of the data shown in Table I.

that could be done to improve the fit?

residuals

To evaluate the quality of a fit, it is very helpful if you plot the *residuals*, which are the differences between the data values and the values calculated by the fitted function, as a function of wavelength. Plot the residuals for the case when you fit with Eq. (1), and for the case when you fit with the extended Eq. (1) that contains one additional oscillator.

Adding even more parameters

Do you see how the data starts "bending down" towards the infrared in Figure 1 ? Can this behavior be modeled at all with Eq. (1)? Can we add one more oscillator so that we can finally obtain a good fit? At which wavelength should this oscillator be, so that it causes the function to bend down to follow the data? Try one last extension to further improve the fit

Too many parameters?

A good question to ask now is when should one stop? Is it really necessary to fit the room temperature data for the refractive index with 6 parameters? Are all 6 parameters determined with enough precision by the data? Is there a way to simplify the dispersion formula so that it contains only 5 parameters but it can still fit the data equally well? What is the advantage of doing so?

Try adding *one more* oscillator and to fit with 8 parameters! What happens, does the fit get better?

Parameter errors

When doing the studies above, try fitting with an algorthm that gives you estimations for the errors in the parameters of the function that are determined by the fit. Levenberg-Marquardt normally does that. What happens to the errors of the parameters when you increase the number of parameters used in the fit?

More...

Once you have determined your preferred "optimum" way to fit this data, try fitting the data at the various temperatures, and analyze what temperature dependence do you get for the parameters of the fitted function. How does the temperature dependence look like when you use too few parameters, too many parameters, or just the right amount?

This is where help from funding agencies or other people is recognized.

[1] W. Sellmeier, "Zur Erklärung der abnormen Farbenfolge im Spectrum einiger Substanzen," Annalen der Physik und Chemie **219**, 272-282 (1871).