

This was originally published in Spectroscopy, 13(11), p. 18-21 (1998)

Chemometrics in Spectroscopy Part 29 - Linearity in Calibration - Act II Scene I

by H. Mark and J. Workman

When we published our recent column "Linearity in Calibration" (1) we didn't quite realize what a firestorm we were going to ignite, although, truth be told, we did not expect everybody to agree with us, either. But if so many actually took the trouble to send their criticisms to us, then there must also be a large "silent majority" out there that are upset, perhaps angry, and almost certainly misunderstanding what we said. We prepared responses to these criticisms, but they became so lengthy that we could not print them all in a single column, therefore they are spread out over the current column and the next few. We will reprint the responses we received at an appropriate point, followed by our comments about them.

Before doing that, however, let's discuss the question of the linearity of spectroscopic data in general, first. There are a number of causes of non-linearity that most chemists and spectroscopists are familiar with. Let us define our terms. When speaking of "linearity" the meaning of the term depends on your point of view, and your interests. An engineer is concerned, perhaps, with the linearity of detector response as a function of incident radiant energy. To a chemist or spectroscopist, the interest is in the linearity of an instrument's readings as a function of the concentration of an analyte in a set of samples. In practice, this is generally interpreted to mean that when measuring a transparent, non-scattering sample, the response of the instrument can be calculated as some constant times the concentration of the analyte (or at least some function of the instrument response can be calculated as a constant times some other function of the concentration). In spectroscopic usage, that is normally interpreted as meaning the condition described theoretically by Beer's Law, i.e., the instrument response function is the negative exponential of the concentration:

$$I = k I_0 e^{-bC} \quad (\text{Eq 1})$$

where:

I = the radiation passing through the sample

k = the multiplying constant

I₀ = the radiation incident on the sample

b = the product of the pathlength and absorbtivity

C = the concentration of the analyte

When other types of samples are measured, the resulting data is usually known to be nonlinear (except possibly in a few special cases) so those measurements are of no interest to us here. Thus, in practice, the invocation of "linearity" implies the assumption that Beer's Law holds, therefore discussions of nonlinearity are essentially about those phenomena that cause departures from Beer's law.

These include:

- 1) Chemical causes
 - a) Hydrogen bonding
 - b) Self-polymerization or condensation
 - c) Interaction with solvent
 - d) Self-interaction
- 2) Instrumental causes
 - a) Non-linear detector
 - b) Non-linear electronics
 - c) Instrument bandwidth broad compared to absorbance band
 - d) Stray light
 - e) Non-collimated radiation
 - f) Excessive signal levels (saturation)

Most chemists and spectroscopists expect that in the absence of these distinct phenomena causing non-linearity, Beer's Law provides an exact description of the relationship between the absorbance and the analyte concentration. Unfortunately the world is not so simple, and Beer's Law never holds exactly, EVEN IN PRINCIPLE. The reason for this arises from thermodynamics.

Optical designers, and specialists in heat transfer calculations in the chemical engineering and mechanical engineering sciences are familiar with the mathematical construct known as The Equation of Radiative Transfer, although most chemists and spectroscopists are not. The Equation of Radiative Transfer states that, disregarding absorbance and scattering, in a lossless optical system:

$$dE = I(\lambda) d\lambda d\omega da dt \quad (\text{Eq 2})$$

where:

dE = the differential energy transferred in differential time dt

$I(\lambda)$ = the optical intensity as a function of wavelength (i.e., the "spectrum")

$d\lambda$ = the differential wavelength increment

$d\omega$ = the differential optical solid angle the beam encompasses

da = the differential area occupied by the beam

For a static (i.e., unvarying with time) system, we can recast equation 2 as:

$$dE/dt = I(\lambda) d\lambda d\omega da \quad (\text{Eq 3})$$

where dE/dt is the power in the beam. The application of these equations to heat transfer problems is obvious, since by knowing the radiation characteristics of a source and the geometry of the system, these equations allow an engineer, by integrating over the differential terms of equation 2 or equation 3, to calculate the amount of energy transferred by electromagnetic radiation from one place to another.

Furthermore, the first law of thermodynamics assures us that dE/dt will be constant anywhere along the optical beam, since any change would require that the energy in the beam be either increased or decreased, which would require that energy would be either created or destroyed, respectively.

Less obviously, perhaps, the second law of thermodynamics assures us that the intensity, $I(\lambda)$, is also constant along the beam, for if this were not the case, then it would be possible to focus all the radiation from a hot body onto a part of itself, increasing the radiation flux onto that portion and raising its temperature of that portion without doing work - a violation of the second law. The constancy of beam energy and intensity has other consequences, some of which are familiar to most of us. If we solve equation 3 for the product $(d\omega da)$ we get:

$$(d\omega da) = dE/dt \times d\lambda/I(\lambda) \quad (\text{Eq 4})$$

All the terms on the right-hand side of equation 4 are constants, therefore for any given wavelength and source characteristics, the product $(d\omega da)$ is a constant, and in an optical system one can be traded off for the other. We are all familiar with this characteristic of optical systems, in the magnification and demagnification of images described by geometric optics. Whenever light is brought to a small focus (i.e., da becomes small) the light converges on the focal point through a large range of angles (i.e., $d\omega$ becomes large) and vice versa. This trade-off of parameters is more obvious to us when seen through the paradigm of geometric optics, but now we see that that is a manifestation of the thermodynamics underlying it all.

We are also familiar with this effect in another context: in the fact that we cannot focus light to an arbitrarily small focal point, but are limited to what we usually call the "diffraction limit" of the radiation in the beam. This effect also comes out of equation 4, since there is a physical (or perhaps a geometrical) limit to $d\omega$: $d\omega$ cannot become arbitrarily large, therefore da cannot become arbitrarily small. Again, we are familiar with this effect by coming across it in another context, but we see that it is another manifestation of the underlying thermodynamic reality.

Getting back to our main line of discussion: we can see from equation 2 (or equation 3) that the differential terms must all have finite values. If any of the terms $d\lambda$, $d\omega$ or da were zero, then zero energy would pass through the system and we could not make any measurements. One thing this tells us, of interest to us as spectroscopists, is that we can never build an instrument with perfect resolution. The mechanistic fundamentals (quantum broadening, Doppler broadening, etc.) have been extensively discussed by one of our fellow columnists(2). This effect also manifests itself in the fact that every technology has an "instrument function" that is convolved with the sample spectrum, and each instrument function is explained by the paradigms of the associated technology, but since "perfect" resolution means that $d\lambda = 0$, we see again that this is another result of the same underlying thermodynamics.

More to the point of our discussion regarding non-linearity, however, is the fact that $d\omega$ cannot be zero. $d\omega$ is related to the concept of "collimation": for a "perfectly collimated" beam, $d\omega = 0$. But as we have just seen, such a beam can transfer zero energy; so just as with $d\lambda$ and da , a perfectly collimated beam has no energy.

Beer's law, on the other hand, is based on the assumption that there is a single pathlength (normally represented by the variable b in the equation $A=abc$) for all rays through the sample. In a real, physical, measurement system, this assumption is always false, because of the fact that $d\omega$ cannot be zero. As figure 1 shows, the actual rays

have pathlengths that range from b (for those rays that travel "straight through", i.e., normal to the sample surfaces) to $b/\cos(\theta_{\max})$ (for the rays at the most extreme angles). We noted this effect above as item 2e in our list of sources of non-linearity, and here we see the reason that there is fundamental limitation. Mechanistically, the non-linearity is caused by the fact that the absorbance for the rays traveling normally = abc , while for the extreme rays it is $abc/\cos(\theta_{\max})$. Thus the non-normal rays suffer higher absorbance than the normal ones do, and the discrepancy (which equals $abc(1-1/\cos(\theta))$) increases with increasing concentration.

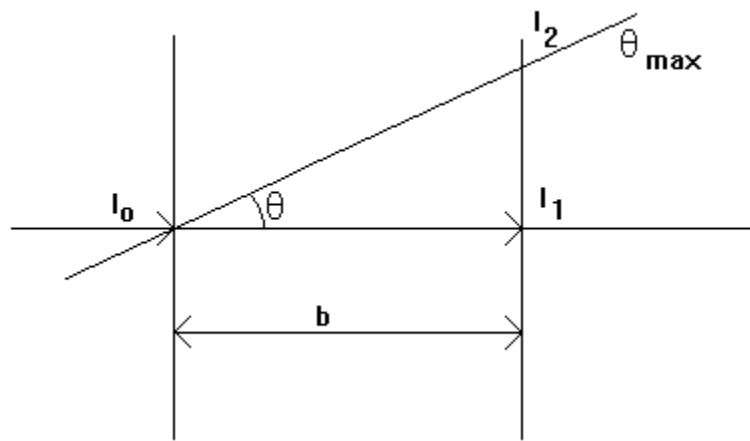
When the medium is completely non-absorbing then the difference in pathlength does not affect the measurement. When the sample has absorbance, however, it is clear that ray I_2 will have its intensity reduced more than ray I_1 , due to the longer pathlength. Thus not all rays are reduced by a different amount and this leads to the non-linearity of the measurement. Mathematically, this can be expressed by noting that the intensity measured when a beam with a finite range of angles passes through a sample is:

$$I = I_0 \int_0^{q_{\max}} e^{-b/\cos(q)} dq \quad (\text{Eq 6})$$

rather than the simpler form shown equation 1 (which, we remind the reader, only holds true for "perfectly collimated" beams, which have zero energy).

In practice, of course, this effect is very small, normally much smaller than any of the other sources of nonlinear behavior, and we are ordinarily safe in ignoring it, and calling Beer's law behavior "linear" in the absence of any of the other known sources of non-linear behavior. However, the point here is that this completes the demonstration of our statement above, that Beer's law never exactly holds IN PRINCIPLE and that as spectroscopists we never ever really work with perfectly linear data.

FIGURE 1



REFERENCES

1. Mark, H., Workman, J.; Spectroscopy; 13 (6), p.19-21 (1998)
2. Ball, D. W.; Spectroscopy; 11 (1), p.29-30 (1996)