

A Simple and Convenient Method of Multiple Linear Regression To Calculate Iodine Molecular Constants

Paul D. Cooper

Department of Chemistry and Biochemistry, George Mason University, Fairfax, Virginia 22030
pcooper6@gmu.edu

The vibronic spectrum of I₂ (Figure 1) is widely used for teaching spectroscopy and the calculation of molecular constants in undergraduate physical chemistry laboratories. The relative ease in which a student can obtain the spectrum experimentally and the reasonably sophisticated molecular properties that can be derived from one data set are the reasons behind the experiment's success.

Prior to the availability of data analysis using computer programs, the iodine experiment was treated using the Birge–Sponer method (1), which is both laborious and not particularly straightforward or intuitive. This method has persisted (2) even with the availability of computer analysis, despite its obvious pedagogical inadequacies (3). A recent article describes the use of MathCad to perform the Birge–Sponer analysis (4). An alternative approach has been the use of least-squares analysis (5–7). Another method (3) uses a second-order polynomial to fit data and solve for the excited-state fundamental vibrational frequency and anharmonicity constant. However, ground-state parameters cannot be calculated by this method.

Over recent years, it has become accepted (5–7) that a least-squares fit to experimental data to directly solve for molecular constants is a superior method, both analytically and pedagogically, to the Birge–Sponer method for analyzing the vibronic spectrum of iodine. However, because of tight budget constraints, I suspect that many teaching laboratories in small-to-moderately sized teaching institutions do not have access to sophisticated data-analysis software that is available to the entire class. Although the analysis described here is not new (5, 6), the primary focus of this article is to describe in detail how to perform this least-squares analysis using software that is accessible to all students. Although many laboratories already perform such an analysis, or one similar to it, a search of the literature found no article explicitly describing the actual mechanics of such an analysis that is of use to both the instructor and student. The method described here also encourages students to use statistical analysis to comment on the reliability of the data and subsequent analysis.

Pedagogically, the method of least-squares analysis is more satisfying as it solves the molecular constants directly, in contrast to the somewhat convoluted Birge–Sponer method. As discussed by others (3), this provides students with a clearer understanding of the molecular constants and their derivation.

Method

The multiple linear least-squares analysis described here should be readily available to all undergraduate physical chemistry

laboratories as it uses the LINEST array function in Microsoft Excel. LINEST is a least-squares function that can calculate the line of best fit to the equation,

$$y = m_1x_1 + m_2x_2 + \dots + b \quad (1)$$

where y is a function of x_1, x_2, \dots ; m_1, m_2, \dots are coefficients corresponding to their respective x value; and b is a constant. In addition, LINEST provides standard errors in m, b , and y values as well as regression statistics such as the coefficient of determination (r^2), the F -statistic, regression sum of squares, and residual sum of squares. I introduce LINEST for analysis of single-variable linear regression at the beginning of the physical chemistry laboratory course. As the students use it regularly in other experiments and are proficient in its use and capabilities, they are comfortable with applying it to a multiple linear-regression analysis and understand the method rather than treating it as a black box that simply provides an answer.

The energy of a quantum vibrational level, v , where $v = 0, 1, 2, \dots$, given by the Morse oscillator approximation is

$$E_{v,\text{Morse}} = \omega_e \left(v + \frac{1}{2} \right) - \omega_e x_e \left(v + \frac{1}{2} \right)^2 \quad (2)$$

where ω_e is the fundamental vibrational frequency and $\omega_e x_e$ is the anharmonicity constant. Therefore, the observed transition frequency, ν , of an optical transition from the ground electronic state to a vibrational level in an excited electronic state is

$$\nu = E_{cl} + E_{v'} - E_{v''} \quad (3a)$$

$$\begin{aligned} \nu = E_{cl} + \omega_e' \left(v' + \frac{1}{2} \right) - \omega_e' x_e' \left(v' + \frac{1}{2} \right)^2 \\ - \omega_e'' \left(v'' + \frac{1}{2} \right) + \omega_e'' x_e'' \left(v'' + \frac{1}{2} \right)^2 \end{aligned} \quad (3b)$$

where double- and single-primed terms correspond to ground and excited-state values, respectively. It is easy to see that eq 3b has a similar form to eq 1 if we treat the squared terms as independent variables to their nonsquared counterparts.

Students start the analysis by the traditional approach of assigning peaks in the absorption spectrum to transitions from the $v'' = 0, 1$, and 2 levels of the ground electronic state to vibrational levels in the excited state using a Delandres table (5, 6). Once assignments have been made, students tabulate the data in Excel as shown in Table 1. The values E_{cl} , ω_e' , $\omega_e' x_e'$, ω_e'' , and $\omega_e'' x_e''$ and their standard errors can be determined in a few keystrokes using LINEST. In agreement with earlier authors, it is my belief that this simple method of analysis is far more pedagogically satisfying than other treatments (3). The students

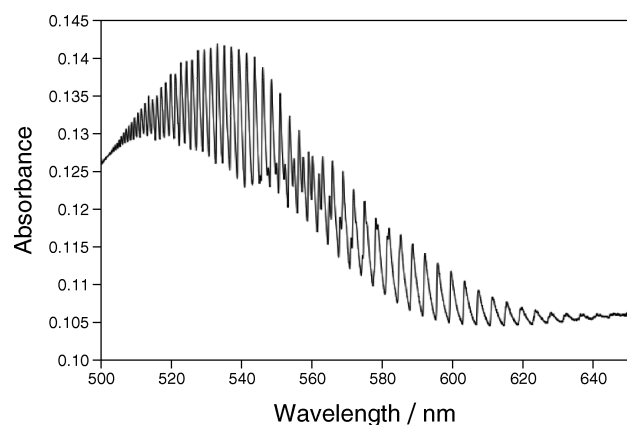


Figure 1. Vibronic spectrum of iodine.

Table 1. An Example of a Spreadsheet Entry in Excel (not all data shown)

v	v'	nm	cm^{-1}	$v + 1/2$	$(v + 1/2)^2$	$v' + 1/2$	$(v' + 1/2)^2$
15	0	574.35	17411	15.5	240.25	0.5	0.25
16	0	571.14	17509	16.5	272.25	0.5	0.25
17	0	568.01	17605	17.5	306.25	0.5	0.25
15	1	581.45	17198	15.5	240.25	1.5	2.25
16	1	578.12	17297	16.5	272.25	1.5	2.25
17	1	574.94	17393	17.5	306.25	1.5	2.25
8	2	615.4	16250	8.5	72.25	2.5	6.25
9	2	611.28	16360	9.5	90.25	2.5	6.25
10	2	607.38	16464	10.5	110.25	2.5	6.25

are not lost in the somewhat laborious and convoluted Birge–Sponer method and the calculations of the fundamental vibrational frequencies and anharmonicity constants are simple once peaks have been assigned. Note that importantly this technique calculates the ground-state and upper-state parameters simultaneously *and* provides the standard error associated with these values.

To use LINEST in Excel, a 5 column \times 5 row area is highlighted using the left mouse button. This area will eventually display the results of the least-squares analysis. The LINEST function can be inserted from the insert function list or can be typed manually as “=LINEST(Known y's, Known x's, Const, Stats)” in the formula bar at the top of the spreadsheet. When selecting data to include in the LINEST array function, all data from both ground and hot bands should be selected. This typically will result in the selection of ~ 50 band heads that will be used for the analysis. Using the left mouse button, select all of the wavenumber values in Table 2 as the Known y's and select all four columns headed $v' + 1/2$, $(v' + 1/2)^2$, $v'' + 1/2$, and $(v'' + 1/2)^2$ as Known x's. For the “const” value, type TRUE as this will allow a nonzero value for the y intercept of the multiple linear least-squares fit. For the “Stats” value, also type TRUE as this will display statistical information regarding the least-squares analysis.

Once these options have been selected, highlight the contents of the formula bar with the left mouse button. An example for the contents of the formula bar is =LINEST(D3:D53,E3:

Table 2. Format of 5 \times 5 LINEST Array Function Output

	A	B	C	D	E
1	$\omega_e''x_e''$	ω_e''	$\omega_e'x_e'$	ω_e'	E_{el}
2	$\pm\omega_e''x_e''$	$\pm\omega_e''$	$\pm\omega_e'x_e'$	$\pm\omega_e'$	$\pm E_{el}$
3	r^2	$s(y)$	#N/A	#N/A	#N/A
4	F	DOF	#N/A	#N/A	#N/A
5	SS_{reg}	SS_{resid}	#N/A	#N/A	#N/A

I53,TRUE,TRUE). Once this is highlighted, on a Mac the Apple and Enter keys are depressed or on a PC the Shift, Ctrl, and Enter keys are depressed. This places braces {...} around the LINEST function. These braces cannot be typed in manually. This prints numerical values to the 5 column \times 5 row array that can then be used for the analysis. For further information regarding the use of LINEST, use the Excel Help feature and search for LINEST for more detailed instructions.

If the data are entered into the spreadsheet as in Table 1, the output from the LINEST array function should be formatted according to Table 2. Row 1 displays the values of the spectroscopic parameters $\omega_e''x_e''$, ω_e'' , $\omega_e'x_e'$, ω_e' , and E_{el} . Row 2 displays the standard error ($\sim 68\%$ confidence interval) in each of the spectroscopic parameters. Cell A3 displays the correlation coefficient of the data set, B3 displays the standard error in y , A4 displays the F -statistic, B4 displays the degrees of freedom of the data set, A5 displays the regression sum of squares, and B5 displays the residual sum of squares.

Results

Typical results for the LINEST analysis of the I_2 vibronic spectrum are shown in Table 3 along with values found in the literature. The LINEST computed values compare favorably with the literature values. The only value that is not calculated well is the ground-state anharmonicity constant, $\omega_e''x_e''$. The deviation from the literature value and the relatively large uncertainty in the value is a result of only having only three vibrational levels ($v'' = 0, 1$, and 2) contribute to determine the anharmonicity of the potential energy curve of the ground state. Although this is an inherent limitation of this analytical method, it can also be used to illustrate the process of least-squares regression and can be discussed in class. The I_2 fluorescence experiment detailed in Garland et al. (6) could be used to provide extra data points containing values of $v'' > 2$, which would provide a better estimate of the ground-state anharmonicity constant.

An additional and easy-to-implement advantage of the LINEST analysis is the ability to test whether an excited-state cubic term ($\omega_e'y_e'(v' + 1/2)^3$) added to eq 3b results in a statistically better fit to the experimental data. The resulting equation is

$$v = E_{el} + \omega_e'(v' + 1/2) - \omega_e'x_e'(v' + 1/2)^2 + \omega_e'y_e'(v' + 1/2)^3 - \omega_e''(v'' + 1/2) + \omega_e''x_e''(v'' + 1/2)^2 \quad (3c)$$

The results of this extra analysis are also given in Table 3. As is shown, with the extra cubic term, better agreement is achieved between LINEST calculated values and the literature values for

Table 3. Comparison of I₂ Spectrum Analysis Methods

Parameter	Least-squares fit without $\omega_e'y_e'$	Least-squares fit with $\omega_e'y_e'$	Williamson (7)	Huber/Herzberg (8)
E_{el}/cm^{-1}	15725 ± 4	15754 ± 4	15757 ± 26	15769.01
$\omega_e''/\text{cm}^{-1}$	217.0 ± 2.4	214.4 ± 1.6	214.36 ± 0.09	214.50
$\omega_e''x_e''/\text{cm}^{-1}$	2.27 ± 0.84	1.34 ± 0.55	0.603 ± 0.008	0.614
ω_e'/cm^{-1}	131.2 ± 0.3	126.9 ± 0.5	128 ± 3	125.69
$\omega_e'x_e'/\text{cm}^{-1}$	0.999 ± 0.005	0.819 ± 0.021	0.79 ± 0.15	0.764
$\omega_e'y_e'/\text{cm}^{-1}$	-	$(-2.3 \pm 0.3) \times 10^{-3}$	$(-6 \pm 3) \times 10^{-3}$	-
r^2	0.9999932	0.9999973	-	-
RSS	256.1	100.9	-	-

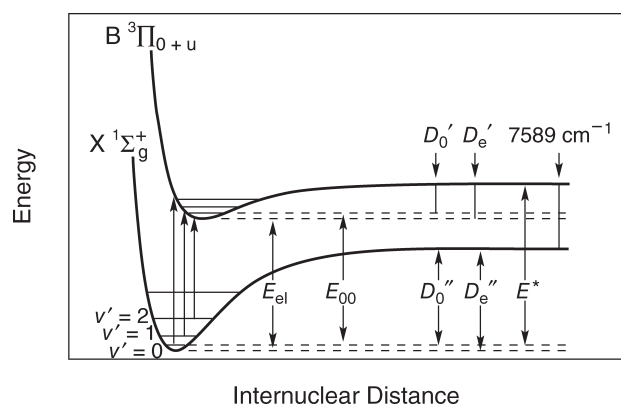


Figure 2. Energy level diagram for X- and B-states of molecular iodine. The positions of vibrational levels are exaggerated for illustrative purposes.

all molecular constants. Importantly, the value of $\omega_e''x_e''$ improves, although the relative error is still large. The correlation coefficient and residual sum of squares also indicate a better statistical fit to the data. Because of the extra term, a 6 column \times 5 row LINEST array needs to be created to perform this analysis.

Additionally, because of the ease of use of LINEST, multiple LINEST calculations can be performed quickly and easily. One particular advantage of this feature is having the ability to rapidly repeat the least-squares analysis after removing some of the data points, in particular some of the broader asymmetric band heads at longer wavelength, to test whether the molecular constants improve.

With the determination of these molecular constants, students can then calculate the dissociation energy, D_e , and bond energy, D_0 (Figure 2). The method briefly described here outlines the procedure for the calculation of D_e and D_0 using eq 3b and the molecular constants determined without the excited-state cubic term. First, students calculate the energy difference, E_{00} , between the $v'' = 0$ and $v' = 0$ states using the values in Table 3 and eq 3b. Second, the energy difference, E^* , between the $v'' = 0$ and the top of the excited-state potential well is calculated. Owing to anharmonicity, the spacing between adjacent vibrational levels approaches zero at higher values of v' . Therefore, students differentiate eq 3b with respect to $(v' + 1/2)$ and set this to 0:

$$0 = \omega_e' - 2\omega_e'x_e'(v' + 1/2) \quad (4)$$

This is the asymptotic limit where the vibrational spacing goes to zero at the point of dissociation. Then, $(v' + 1/2)_{\max}$ is expressed

Table 4. Comparison of Bond and Dissociation Energies of I₂

	This Work	Williamson (7)
D_0''/cm^{-1}	12333	12440
D_e''/cm^{-1}	12442	12547
D_0'/cm^{-1}	4239	4270
D_e'/cm^{-1}	4304	4330

in terms of ω_e' and $\omega_e'x_e'$ and substituted back into eq 3b to calculate E^* .

$$E^* = E_{el} + \omega_e' \left(\frac{\omega_e'}{2\omega_e'x_e'} \right) - \omega_e'x_e' \left(\frac{\omega_e'}{2\omega_e'x_e'} \right)^2 - \omega_e'' \left(v'' + \frac{1}{2} \right) + \omega_e''x_e'' \left(v'' + \frac{1}{2} \right)^2 \quad (5)$$

The excited-state bond energy, D_0' , the difference in energy from the $v' = 0$ level to the top of the excited-state potential well, can now be calculated from E^* and E_{00} . The excited-state dissociation energy, D_e' , the difference in energy in the excited-state from the bottom of the potential well to the dissociation limit can be calculated from D_0' and the excited-state zero-point energy, ZPE' . The latter is calculated by substituting $v' = 0$ into eq 6.

$$ZPE' = \omega_e' \left(v' + \frac{1}{2} \right) + \omega_e'x_e' \left(v' + \frac{1}{2} \right)^2 \quad \text{where } v' = 0 \quad (6)$$

The ground-state bond energy, D_0'' , can be calculated from E^* and knowing that the energy difference between the dissociation point for upper and ground states is 7589 cm^{-1} (5). The ground-state dissociation energy, D_e'' , can be calculated from D_0'' and the ground-state zero-point energy, which can be calculated by substituting $v'' = 0$ into eq 6:

$$ZPE'' = \omega_e'' \left(v'' + \frac{1}{2} \right) + \omega_e''x_e'' \left(v'' + \frac{1}{2} \right)^2 \quad \text{where } v'' = 0 \quad (7)$$

Values for D_0'' , D_e'' , D_0' , D_e' are shown in Table 4 and compared with literature values. Students can then plot the potential-energy surfaces, $U(r)$, of the ground and excited states of I₂ using the Morse function

$$U(r) = D_e[1 - e^{-\beta(r-r_e)}]^2$$

$$\beta = \omega_e \sqrt{\frac{2\pi^2\mu c}{D_e h}} \quad (8)$$

where r is the bond length, c is the speed of light, h is Planck's constant, and μ is the reduced mass.

Summary

The method described above is particularly advantageous for a number of reasons. First, the method of least-squares regression to solve eq 3b is pedagogically more satisfying (3) and allows students to understand better the mathematical approach to solving for molecular constants than the more mathematically obscure, yet well-established, Birge–Sponer method. Second, the use of LINEST provides a multitude of advantages that include widespread availability to students, ease of use, and output of standard errors that can be incorporated into the analysis. Third, if the use of LINEST is introduced early into a physical chemistry laboratory program, it can be used to solve simple linear plots that are common in such courses, in addition to the rovibrational spectrum of HCl (6) and other experiments that require multivariable analysis.

Literature Cited

1. Stafford, F. E. *J. Chem. Educ.* **1962**, *39*, 626–629.
2. Lessinger, L. J. *J. Chem. Educ.* **1994**, *71*, 388–391.
3. Pursell, C. J.; Doezeema, L. J. *J. Chem. Educ.* **1999**, *76*, 839–841.
4. Long, G.; Sauder, D.; Shalhoub, G. M.; Stout, R.; Towns, M. H.; Zielinski, T. J. *J. Chem. Educ.* **1999**, *76*, 841–847.
5. McNaught, I. J. *J. Chem. Educ.* **1980**, *57*, 101–105.
6. Garland, C. W.; Nibler, J. W.; Shoemaker, D. P. *Experiments in Physical Chemistry*, 7th ed.; McGraw Hill: New York, 2003.
7. Williamson, J. C. *J. Chem. Educ.* **2007**, *84*, 1355–1359.
8. Huber, K. P.; Herzberg, G. Constants of Diatomic Molecules. In *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*; Linstrom, P. J., Mallard, W. G., Eds.; National Institute of Standards and Technology: Gaithersburg, MD 20899 (<http://webbook.nist.gov>).