CROATICA CHEMICA ACTA
CCACAA **79** (4) 559–571 (2006)

ISSN-0011-1643
CCA-3125
Original Scientific Paper

A Periodicity-Sensitive Vector Index for Small Molecules

Ray Hefferlin* and Ken Luk

Physics Department, Southern Adventist University, Collegedale, Tennessee 37315, United States of America

RECEIVED OCTOBER 18, 2005; REVISED FEBRUARY 22, 2006; ACCEPTED APRIL 27, 2006

Keywords:
molecular vector index
XY_n molecules
molecular periodicity
second periodicity
matrix data transformations
inverse least-squares

In the vastness of molecular space, there are many series X, XY, …, XY_n…XY_N, where N lies between 3 and say 10, whose data for a given property and phase are approximately linear with respect to n. A vectorial representation of the tabulated data in a series and a vector index to describe the series have been developed. The authors started with X as a metallic atom and with the property as heat of atomization, and showed that the vector index manifested periodicity. Then they moved to cases where X itself is a molecule and where the properties are enthalpies of formation, entropy, retention index, hydrophobicity, and boiling point. The vector index is a two-dimensional vector whose upper element describes the value of the property for the atom or molecule X and whose lower element describes the abscissa difference of any two members of the series after the data have been fitted, in least-squares fashion, to a standard, linear with n, series A, AL, … AL_n, … AL_N. Matrices can transform the data vectors of any series of species whatsoever to any other series of the same dimensionality. Matrices can also transform the vector index for property data of any approximately linear series, in any phase, to the vector index for any other approximately linear series.

INTRODUCTION

Background

A large group of Russian chemists at the Saint Petersburg University (LSU, earlier Leningrad State University) devoted decades to compiling molecular databases for small molecules, in graphical form for maximum impact, shelved in a manner consistent with chemical periodicity.[1–6] One of the motivations for this effort was a vision that it would be possible, eventually, to go from the entries in one database, *e.g.,* group-2 halides, to those in another, *e.g.,* transition-metal oxides.[7] References 3 and 4 contain graphs for the enthalpy of atomization ΔH_a for molecules formed from group 1-oxides; Mg and Ca oxides; groups-4 to 8 oxides; group-13 hydrides and halides; group-14 and 15 hydrides and oxides; group-16 hydrides, oxides, and sulfides; group-17 hydrides, oxides, and fluorides; and group-18 fluorides. Graphs have the oxidation state of the central atom on the x axis (abscissa) and ΔH_a on the ordinate. References also contain many plots for the standard enthalpy of formation ΔH_f° (298.15 K). (Some graphs have as X a molecule, such as N_2O_3, and in these cases both values are normalized to one »metal« atom, in this instance divided by two.) These databases are the first of four starting points for the work reported in this paper.

Inspection of the database graphs for either of these two properties reveals that they are qualitatively similar, that they echo periodic behaviors of the central atoms, and that they suggest linearity as a function of the num-

* Author to whom correspondence should be addressed. (E-mail: hefferln@southern.edu)

ber of added ligand atoms, n. Many of these graphs were replotted in various ways during our study to explore these phenomena more thoroughly (and at the same time to employ more recent data). The most useful representation is for the x axis to be the total number of atoms in the molecules, $1 + n$. The suggestion of linearity is consistent with other findings that the binding energy per ligand (*e.g.,* H attached to a carbon atom) is roughly constant for many molecules.[8,9] The linearity is not perfect because the ligands interact (the »endoeffect«[7]) and also affect the central atom. The approximate linearity of these graphs is the second starting point for this essay.

The periodicity of properties of diatomic molecules has been studied so thoroughly that a periodic system has been constructed for them. Actually there are several good systems (just as there are various good charts of the elements) – additive,[10] outer-matrix product,[11] and group-dynamic.[12] Each of these systems can be generalized to larger molecules,[12–14] though demonstration of their faithfulness to reality is demonstrable only by averaging the properties for triatomic [13] and four-atom[15] molecules (***N*** = 3 and 4) and seems virtually impossible beyond that. By abandoning the obsession to treat all molecules with ***N*** atoms and instead selecting series of molecules for which data are available, one can explore general trends deeper in molecular space than ***N*** = 4 (but not so deep as to where biological species are found). This exploration is the third starting point for this work.

Quantum computation cannot yet produce a *global* demonstration of molecular periodicity – computation, like experiment, yields numerical values for just one or for a few (*e.g.*, isoelectronic) molecules at a time; these have to be assembled to demonstrate that periodicity exists in molecular data. At the moment, periodicity remains a fundamental reality of chemistry.[16,17]

Decades of research in discrete mathematical chemistry have resulted in the formulation of hundreds of topological indices and combinations of indices, documented in an immense literature, that describe molecular graphs.[18–21] They are applicable primarily to organic molecules, and some of them have been retrofitted for atoms other than carbon (heteroatoms). The excellent J_{het} index of Balaban[22,23] has been additionally fitted so as to include periodicity; however, the operative parameter is not available for all atoms. These indices are the final starting point for this paper.

Goals

1. To define, for series of small molecules XY_n with property ΔH_a, an index that characterizes the series;
2. To show that the method is valid for more complex species and for other properties;
3. To transform the data, and the vector indices, for some property for one set of species to any other set (thus fulfilling the vision of the group at LSU).

Definitions

We use the word »species« to mean atom(s), molecule(s), or both. By a »series« we shall mean a central atom and the molecules formed by the bonding of its ligands; by »ligands« we mean *any* species attached to a central atom or molecule. The end of a data series is when the maximum number of ligands, N, are bonded. We do not consider series with only two species and in reality we like N to be as large as possible (such as eight, for oxides of iron).

THEORY

Definitions

We begin with the ΔH_a, which describes the reaction:

$$X + nY \rightarrow XY_n + \Delta H_a \qquad (1)$$

$\Delta H_a = 0$ if $n = 0$, *i.e.*, for a lone atom, and is negative for stable molecules $n > 1$.

Let there be one standard atom, A, whose ΔH_a is of course zero. Let there be n ligands L. Let the data for ΔH_a be plotted on a graph with an x axis enumerating $n + 1$ (total number of atoms). The point for A is located at $(x, y) = (1,0)$ and the bonding of each successive L increases x by 1 and lowers ΔH_a by an increment of 1000; the abscissae for AL_n will be $x(AL_n) = (n + 1)$ and the ordinates will be $(-1000)\ n$. Now let data for ΔH_a of any real-world system of interest, consisting of a central atom X with N ligands Y, be plotted on the same graph. The point for X is located at $(x, y) = (1,0)$ and the bonding of each successive Y increases x by 1 and lowers ΔH_a by increments; the abscissae for XY_n will be $x(XY_n) = (n + 1)$ and the ordinates will be $\Delta H_a(XY_n) < 0$

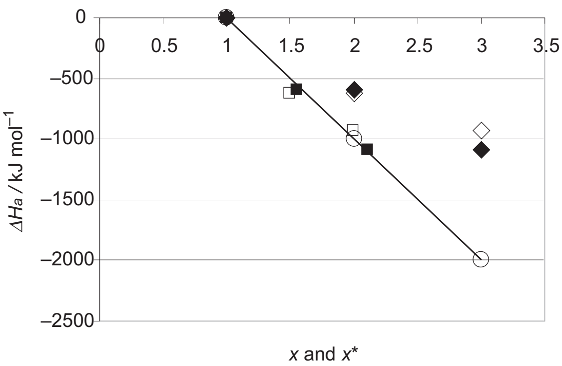

Figure 1. ◆ and ◇: raw data ΔH_a for N, NO, and NO_2, and for P, PO, and PO_2 are plotted against x, the total number of atoms in the molecules. ○: data for the standard atom A and standard molecules AL and AL_2. ■ and □: the raw data fitted, by inverse least-squares, to the line while keeping the two differences of the three abscissae equal. The original abscissae (diamonds) are x; the fitted abscissae (squares) are x^*, and if the points were forced to be exactly on the line for A to AL_2, then their abscissae would be x^o.

(Figure 1). This figure shows a series with only three members; it is with such series that we begin the mathematical articulation that follows.

Imagine that we slide the data points for the real-world molecules XY and XY_2 horizontally so that they fit as closely as possible to the line for the standard series A, AL, and AL_2, while at the same time keeping the differences in their abscissae equal to *each other* (but not necessarily equal to 1 as they were when originally plotted). Using x^* represent the new abscissae; the difference ($x^*(XY_2) - x^*(XY)$) will be equal to ($x^*(XY) - x^*(X)$, or equal to ($x^*(XY) - 1$) (Figure 1). In actuality, this mapping amounts to a manual, horizontal, inverse least-squares procedure – fitting the points to a *defined straight line*. The equal abscissa differences $x^*(XL_2) - x^*(XL)$ and $x^*(XL) - 1$ will be used below to characterize the series.

Before laying out the derivation, we should answer the question: »Why not characterize the data for the series by means of the slopes of their least-squares trend lines?«. The response: simple trend lines will not serve because they will miss the atom points at (1,0). While a mathematical expression for a trend line pivoted at (1,0) can certainly be derived; the derivation is not much simpler than that to be presented here.

Mathematical articulation

Now given X, XY, and XY_2, it is required that:

$$x^*(XY_2) - x^*(XY) = x^*(XY) - 1 \tag{2}$$

Next, it is required that the points x^* for XY and XY_2 (not for X, which is fixed at its original location) be fitted in a least-squares fashion to the line for the standard species:

$$\delta\{[x^*(XY_2) - x^o(XY_2)]^2 + [x^*(XY) - x^o(XY)]^2\}^{1/2} = 0 \tag{3}$$

where $x^o(XY)$ and $x^o(XY_2)$ are the abscissae corresponding to the real-world species if they were to fall, when slid horizontally, exactly onto the line for the standard species. Due to the way in which the locations of A, AL, and AL_2 were defined, it follows that:

$$x^o(XY) = |\Delta H_a(XY)| \ / \ 1000 + 1 \tag{4}$$

$$x^o(XY_2) = |\Delta H_a(XY_2)| \ / \ 1000 + 1 \tag{5}$$

From Equations (2) to (4), we have:

$$\delta\{[x^*(XY_2) - [|\Delta H_a(XY_2) \ / \ 1000 + 1]]^2 + [x^*(XY) - [|\Delta H_a(XY)| \ / \ 1000 + 1]]^2\}^{1/2} = 0 \tag{6}$$

When $x^*(XY_2)$ is expressed in terms of $x^*(XY)$ in Eq. (1) and then substituted into Eq. (3), the minimum condition is expressed in terms of only one variable, $x^*(XY)$. This variable can now be determined by taking the partial derivative of the argument of Eq. (6) with respect to $x^*(XY)$, setting it equal to zero, and solving for $x^*(XY)$. The result may then be substituted into Eq. (2) to find $x^*(XY_2)$.

The derived equations are as follows:

$$x^*(XY) = 1 + (1/5)[|\Delta H_a(XY)| \ / \ 1000] + (2/5)[|D(XY_2)| \ / \ 1000] \tag{7}$$

$$x^*(XY_2) = 1 + (2/5)[|\Delta H_a(XY)| \ / \ 1000] + (4/5)[D(XY_2) \ / \ 1000] \tag{8}$$

Now consider a series ending in XY_3. It is necessary to supplement Equation (2) with an expression that equates the abscissa difference between the points for XY_3 and XY_2 to the abscissa difference between the points for XY_2 and XY:

$$x^*(XY_3) - x^*(XY_2) = x^*(XY_2) - x^*(XY) \tag{9}$$

The substitutions and differentiation then result in the following:

$$x^*(XY) = 1 + (1/14)[|\Delta H_a(XY)| \ / \ 1000] + (2/14)[|\Delta H_a(XY_2)| \ / \ 1000] + (3/14)[|D(XY_3)| \ / \ 1000] \tag{10}$$

$$x^*(XY_2) = 1 + (2/14)[|\Delta H_a(XY)| \ / \ 1000] + (4/14)[|\Delta H_a(XY_2)| \ / \ 1000] + (6/14)[|\Delta H_a(XY_3)| \ / \ 1000] \tag{11}$$

$$x^*(XY_3) = 1 + (3/14)[|\Delta H_a(XY)| \ / \ 1000] + (6/5)[|\Delta H_a(XY_2)| \ / \ 1000] + (9/14)[|\Delta H_a(XY_3)| \ / \ 1000] \tag{12}$$

For the series ending in XY_4, a similar procedure results in:

$$x^*(XY) = 1 + (1/30)[|\Delta H_a(XY)| \ / \ 1000] + (2/30)[|\Delta H_a(XY_2)| \ / \ 1000] + (3/30)[|\Delta H_a(XY_3)| \ / \ 1000] + (4/30)[|\Delta H_a(XY_4)| \ / \ 1000] \tag{13}$$

$$x^*(XY_2) = 1 + (2/30)[|\Delta H_a(XY)| \ / \ 1000] + (4/30)[|\Delta H_a(XY_2)| \ / \ 1000] +$$

$$(6/30)[|\Delta H_a(XY_3)| \ / \ 1000] +$$
$$(4/30)[|\Delta H_a(XY_4)| \ / \ 1000] \tag{14}$$

$$x^*(XY_3) = 1 + (3/30)[|\Delta H_a(XY)| \ / \ 1000] + (6/30)[|\Delta H_a(XY_2)| \ / \ 1000] + (9/30)[|\Delta H_a(XY_3)| \ / \ 1000] + (12/30)[|\Delta H_a(XY_4)| \ / \ 1000] \tag{15}$$

$$x^*(XY_4) = 1 + (4/30)[|\Delta H_a(XY)| \ / \ 1000] + (8/30)[|\Delta H_a(XY_2)| \ / \ 1000] + (12/30)[|\Delta H_a(XY_3)| \ / \ 1000] + (16/30)[|\Delta H_a(XY_4)| \ / \ 1000] \tag{16}$$

(Recall that in every case, $x^*(X) = 1$.) Clear trends are visible among the numerators of the coefficients in these expressions. There is also lawful behavior for the denominators: for the general case:

$$x^*(XY_n) = 1 + (\alpha_N)^{-1} \sum_{\nu=1}^{N} \nu \cdot n \cdot [|\Delta H_a(XY_\nu)| = / \, 1000] \tag{17}$$

where ν is a dummy index, N is the maximum number of ligands, and α_N is:

$$\alpha_N = (1/3)N^3 + (1/2)N^2 + (1/6)N \tag{18}$$

A check of Eq. (17) is provided by considering just A and AL_n and substituting their values $\Delta H_a(AL_n)$ for $\Delta H_a(XY_n)$ in Eq. (17). The result corresponds to the definition given above:

$$x^*(XY_n) = 1 + n = x^*(AL_n) \tag{19}$$

Formulations for other properties or more general species

We have discussed a case where the real-world data have their first point X at $(x,y) = (1,0)$ and where subsequent points lie approximately on a line with a negative slope. This is only the first of several different approximately linear cases that will be considered:

1. Atom X at (1,0) and subsequent points with negative ordinates;
2. Atom X at $(1,y)$ and subsequent points with negative ordinates (which may mean that the points straddle the x axis if $y > 0$); this will happen for ΔH_a if we generalize the central atom X so that it becomes a molecule, or will happen if we take up another property such as the standard enthalpy of formation, $\Delta H_f^o(298.15\text{ K})$;
3. X at (1,0) and subsequent points with positive ordinates;
4. X at $(1,y)$ and subsequent points with positive ordinates (which may mean that the points straddle the x axis if $y < 0$); this can happen for standard entropy, S^o_{298}.

In addition to these cases, there may be a situation where one or more missing molecules invite interpolative prediction.

*The Archive Data Vector **U** and the Vector Index **V***

We may archive the original data in a vector characterizing the series. Its dimension equals the number of members in the series and each element consists of one tabulated datum and its tabulated random error Δ'. The *identifier* for the vector provides information on the central atom, the ligand, the *total* number of objects in the series (1 plus the number of ligands N), and the data property being studied:

$$\boldsymbol{U}(X,Y,(N{+}1),\Delta H_a) = \begin{pmatrix} x(X) \pm \Delta' x(X) \\ x(XY) \pm \Delta' x(XY) \\ \cdots \\ x(XY_n) \pm \Delta' x(XY_n) \end{pmatrix} \tag{20}$$

Finally, we define the index characterizing ΔH_a for each series. This is the *vector index.*

$$\boldsymbol{V}(X,Y,(N+1),\Delta H_a) = \begin{pmatrix} \Delta H_a(X) \\ |x^*(XY_n) - x^*(XY_{n-1})| \exp \Delta x^* \end{pmatrix} \tag{21}$$

The upper element of the vector, $\boldsymbol{V}_1$, represents the property value of the first point – for Case 1 it is zero. The lower element $\boldsymbol{V}_2$ gives any one of the abscissa difference(s) $x^*(XY_{n+1}) - x^*(XY_n)$.

For example, consider N through NO_2 and P through PO_2 (Figure 1).

$$\boldsymbol{V}(N,O,3,\Delta H_a) = \begin{pmatrix} 0 \\ 0.494 \exp 19.83 \end{pmatrix} \tag{22}$$

$$\boldsymbol{V}(P,O,3,\Delta H_a) = \begin{pmatrix} 0 \\ 0.554 \exp 5.05 \end{pmatrix} \tag{23}$$

Case 2, it will be shown, differs in that the first point no longer lies on the x-axis – and so all the data are raised or lowered so that the first point lies at (1,0). $\boldsymbol{V}_1$ is no longer a zero but is the y value of the first point before the normalization. In Case 3, once again the original point is at (1,0), so in order to revert to Case 1 the data are simply reflected through the x axis by multiplying each y value by –1. The vector represents the flipping of the data by carrying an asterisk on the value of $\boldsymbol{V}_2$.

$$V(\mathrm{X},\mathrm{Y},(N+1),p) = \begin{pmatrix} p(\mathrm{X}) \\ |x^*(\mathrm{XY}_n) - x^*(\mathrm{XY}_{n-1})|^* \exp\Delta x^* \end{pmatrix} \quad (24)$$

where p represents the datum for ΔH_a (or any other property) of X. (It is tempting to use the imaginary indicator $\boldsymbol{i}$ instead of the asterisk, but in later derivations this choice would create considerable confusion.) Case 4 is a blend of Cases 2 and 3 because the data need to be normalized as well as reflected through the origin.

Error Notation

The exponential notation in Equations (20) and (21) conforms to an approximate rule for error propagation, which states that when quantities are multiplied or divided, then the percent errors are added to obtain the percent error of the product or quotient. With one minor proviso, this rule makes it possible to write:

$$(|x_i - x_i^*|) \exp\Delta x_i^* \times (|x_j - x_j^*|) \exp\Delta x_j^* =$$

$$(|x_i - x_i^*| \times |x_j - x_j^*|) \exp(\Delta x_i^* + \Delta x_j^*) \quad (25)$$

where the Δx_i^* and Δx_j^* are *percent* errors. The proviso is that in division the exponents are still added.

Transformation Matrices

Now consider transformations from one vector to another. Any set of data, archived in a vector $\boldsymbol{U}$, can be transformed into another set for a different series, property, and phase – the only limitation being if one series has a different N than the other. Vector indices can also be transformed – with no limitation. Suppose that two different series, with central atoms X and X', and with N ligands Y and N' ligands Y', have different properties p and p'. Then the matrix that leads from vector index $\boldsymbol{V}$ to $\boldsymbol{V}'$ *is* $\boldsymbol{M}\{\boldsymbol{V}[\mathrm{X},\mathrm{Y},(N+1),p)] \rightarrow \boldsymbol{V}'[\mathrm{X}',\mathrm{Y}',(N'+1),p']\}$. If both properties conform to Case 1, then the determination of the 2×2 transformation matrix is made trivial by specifying it to be diagonal. A less simple situation is when the matrix transforms an index conforming to Case 2 to another index conforming to Case 1. An example is when the vector index for the series beginning with N_2O_3 [Eq. (44), below] is transformed into the vector index for the series with central atom P [Eq. (39), below], the property in both cases being ΔH_a:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \times \begin{pmatrix} -1590.79 \exp 0.0377 \\ 0.288 \exp 7.29 \end{pmatrix} = \begin{pmatrix} 0 \\ 0.554 \exp 5.13 \end{pmatrix} \quad (26)$$

Setting the matrix elements $a_{11} = 1$ (as usual) and $a_{21} = 0$ gives:

$$(-1591 \exp 0.0377) + a_{12}(0.554 \exp 5.13) = 0 \quad (27)$$

$$a_{22}(0.288 \exp 7.29) = 0.554 \exp 5.13 \quad (28)$$

Thus the transformation matrix is:

$$\boldsymbol{M}[\boldsymbol{V}(\mathrm{N_2O_3},\mathrm{O},3,\Delta H_a) \rightarrow \boldsymbol{V}(\mathrm{P},\mathrm{O},3),\Delta H_a] =$$

$$\begin{pmatrix} 1 & 2743 \exp 5.14 \\ 0 & 0.572 \exp 12.06 \end{pmatrix} \quad (29)$$

If the matrix were to transform an index conforming to Case 4 to another conforming to Case 1, then the vector left-multiplied in Equation (26) would have an asterisk on the magnitude in $\boldsymbol{V}_2$; consequently, a_{22} in Eqs. (28) and (29) would have asterisks. If the matrix transforms an index corresponding to Case 2 to another index corresponding to Case 3, then the vector on the right hand side of Eq. (26) will have an asterisk on the magnitude of $\boldsymbol{V}_2$; therefore, a_{12} and a_{22} in Eqs. (27) to (29) would have asterisks. If the matrix transforms an index fitting Case 4 to another index fitting Case 3, then the $\boldsymbol{V}_2$ of both vectors will have an asterisk and hence a_{12} will have one asterisks and a_{22} will have two. This last instance shows the confusion that would result if, as mentioned under Eq. (24), reflections through the x axis were flagged by the use of imaginary numbers.

There are additional situations were neither vector belongs to Case 1 or 3. We provide a symbolic example:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \times \begin{pmatrix} \mathbf{U}_1 \\ \mathbf{U}_2 \end{pmatrix} = \begin{pmatrix} \mathbf{U}_1' \\ \mathbf{U}_2' \end{pmatrix} \quad (30)$$

Setting the matrix elements $a_{11} = 1$ and $a_{21} = 0$ as before gives:

$$a_{12} = (\boldsymbol{U}_1' - \boldsymbol{U}_1) / \boldsymbol{U}_2 \quad (31)$$

$$a_{22} = (\boldsymbol{U}_2' / \boldsymbol{U}_2) \quad (32)$$

If in this example only $\boldsymbol{U}_2$ has an asterisk (Case 2 to Case 4), then both a_{12} and a_{22} in Eqs. (31) and (32) will be starred; if only $\boldsymbol{U}_2'$ has an asterisk (Case 4 to Case 2) then only a_{22} will have one; and if both vectors have an asterisk (Case 4 to Case 4), then a_{12} would have one asterisk and a_{22} will have two.

Missing Points

It may easily happen that a molecule in the series is missing. In that case, the derivations given above differ. To illustrate the difference with a specific example, we take the special case where data are known for X, XY_2, and XY_3 but not for XY. Eq. (2) must be changed to provide an initial step of two units rather than one:

$$2(x^*(\mathrm{XY_3}) - x^*(\mathrm{XY_2})) = x^*(\mathrm{XY_2}) - 1 \quad (33)$$

When this is supplemented by Eq. (9), a new series of equations follows: Eq. (10) disappears and Eqs. (11) and (12) are replaced by:

$$x^*(XY_2) = 1 + 4/13(|D(XY_2)| \ / \ 1000) + 6/13(|D(XY_3)| \ / \ 1000) \quad (34)$$

$$x^*(XY_3) = 1 + 6/13(|D(XY_2)| \ / \ 1000) + 9/13(|D(XY_3)| \ / \ 1000) \quad (35)$$

Note that in Equation (34) one would expect a term $2/13(|D(XY)| \ / \ 1000)$ to follow the »1«, but it does not appear because of the absence of data for XY. A similar comment applies to Eq. (35).

RESULTS

Case 1: X at (1,0) and Subsequent Points with Negative Ordinates

The data[24] for ΔH_a of (gaseous) nitrogen and phosphorus oxides are plotted in Figure 1. The data-archive vectors are:

$$\boldsymbol{U}(\mathrm{N,O},3,\Delta H_a) = \begin{pmatrix} 0 \\ -626.84 \pm 0.2 \\ -927.384 \pm 0.6 \end{pmatrix} \quad (36)$$

$$\boldsymbol{U}(\mathrm{P,O},3,\Delta H_a) = \begin{pmatrix} 0 \\ -590 \pm 2 \\ -1087.13 \pm 6 \end{pmatrix} \quad (37)$$

The errors for ΔH_a are given[24] by letters, which mean the following (in kJ/mol): A, ≤ 0.1; B, ≤ 0.3; C, ≤ 1; D, ≤ 3; E, ≤ 10; F, ≤ 30; and G, > 30. We chose to encompass these measures quantitatively as follows: A: 0.1; B: 0.2; C: 0.6; D: 2; E: 6; F: 20; G: 60 (in kJ/mol). A transformation from one of these data sets to the other can be accomplished easily with a diagonal matrix. Since both series have the same number of species, the transformation will be complete.

The vector indices are from Eqs. (7) and (8):

$$\boldsymbol{V}(\mathrm{N,O},3,\Delta H_a) = \begin{pmatrix} 0 \\ 0.494 \exp 19.99 \end{pmatrix} \quad (38)$$

$$\boldsymbol{V}(\mathrm{P,O},3,\Delta H_a) = \begin{pmatrix} 0 \\ 0.554 \exp 5.13 \end{pmatrix} \quad (39)$$

The zeros in $\boldsymbol{V}_1$ signify that ΔH_a for *gaseous* metal atoms are by definition zero. The exponents in $\boldsymbol{V}_2$ are the average of the absolute differences between x^* and x^o, for NO and NO_2 and for PO and PO_2, expressed as percent of the abscissa differences $(x^*(XY_2) - x^*(XY))$, being equal to $(x^*(XY) - 1)$.

Central atom	F	Cl	Br	I
Be	$\begin{pmatrix} 0 \\ -575 \pm 6 \\ -1271 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -385 \pm 6 \\ -921 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -300 \pm 20 \\ -772 \pm 20 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -220 \pm 20 \\ -595 \pm 6 \end{pmatrix}$
Mg	$\begin{pmatrix} 0 \\ -455 \pm 6 \\ -1036 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -320 \pm 6 \\ -783 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -250 \pm 20 \\ -669 \pm 20 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -190 \pm 20 \\ -521 \pm 20 \end{pmatrix}$
Ca	$\begin{pmatrix} 0 \\ -530 \pm 6 \\ -1121 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -400 \pm 6 \\ -902 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -318 \pm 6 \\ -687 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -270 \pm 6 \\ -651 \pm 6 \end{pmatrix}$
Sr	$\begin{pmatrix} 0 \\ -540 \pm 6 \\ -1098 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -405 \pm 6 \\ -886 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -335 \pm 6 \\ -784 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -271 \pm 6 \\ -644 \pm 6 \end{pmatrix}$
Ba	$\begin{pmatrix} 0 \\ -580 \pm 6 \\ -1131 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -440 \pm 6 \\ -906 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -370 \pm 6 \\ -817 \pm 6 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -302 \pm 6 \\ -679 \pm 6 \end{pmatrix}$

Ligand atom

Figure 2. Archive vectors for ΔH_a data of alkali-earth halogens.

Central atom	F	Cl	Br	I
Be	$\begin{pmatrix} 0 \\ 0.623 \exp 5.78 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.446 \exp 10.09 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.369 \exp 14.09 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.282 \exp 16.31 \end{pmatrix}$
Mg	$\begin{pmatrix} 0 \\ 0.505 \exp 7.52 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.377 \exp 11.41 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.318 \exp 16.04 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.246 \exp 17.07 \end{pmatrix}$
Ca	$\begin{pmatrix} 0 \\ 0.554 \exp 3.25 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.441 \exp 6.80 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.378 \exp 11.90 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.314 \exp 10.51 \end{pmatrix}$
Sr	$\begin{pmatrix} 0 \\ 0.547 \exp 0.91 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.435 \exp 5.29 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.380 \exp 8.95 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.312 \exp 9.94 \end{pmatrix}$
Ba	$\begin{pmatrix} 0 \\ 0.568 \exp 1.58 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.450 \exp 1.78 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.402 \exp 4.98 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.332 \exp 6.81 \end{pmatrix}$

Ligand atom

Figure 3. Vector indices for ΔH_a data of alkali-earth halogens.

Now we take up group-2 halides. It is not necessary to show graphs for all the 20 series, as they resemble the two series shown in Figure 1. The slight non-linearity of the points has a negative curvature, suggesting that the metal atom seeks fulfillment in the octet rule by adding (in these series) the second halide. Figures 2 and 3 give

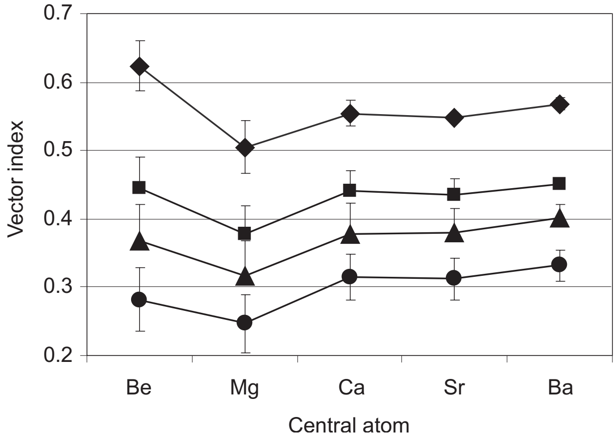

Figure 4. Variation of the magnitudes of V_2 of the vector indices in Figure 3, showing how ΔH_a data of halides of group-2 atoms vary as the central atoms are changed. The sequence of patterns from top (fluorides) to bottom (iodides) is a manifestation of chemical periodicity. Larger gaps between the patterns for fluorides and chlorides, and between bromides to iodides, are a manifestation of secondary periodicity. Within any given pattern, the changes from beryllium to magnesium and from calcium to strontium have less positive slopes than the other changes, which is another manifestation of secondary periodicity. The bars are percentage errors; lower bars are not shown the two central patterns.

the archive vectors $\boldsymbol{U}$ and the vector indices $\boldsymbol{V}$ for the whole set. Figure 4 shows that $\boldsymbol{V}_2$ are periodic for both the central and ligand atoms; secondary periodicities are manifested by the different slopes going from left to right and by the differences in ordinates of the »curves.« Figures 2 and 3 may be thought of as graphs containing single-column matrices, and are thus the converse of the graphical matrices of Nikolić, Miličević, and Trinajstić.[25]

A matrix may be constructed to transform any one of these vector indices into any other. The one closest to being a unit matrix is:

$$\boldsymbol{M}\{[\boldsymbol{V}(\mathrm{Mg,Cl,3,}\Delta H_a(\mathrm{kJ/mol})]\rightarrow\boldsymbol{V}(\mathrm{Ca,Br,3,}\Delta H_a(\mathrm{kJ/mol})]\}$$

$$= \begin{pmatrix} 1 & 0 \\ 0 & 1.00265\exp 0.232 \end{pmatrix} \tag{40}$$

$\boldsymbol{M}[(\mathrm{Sr,I,3} \rightarrow \mathrm{Ca,I,3}),\Delta H_a]$ and $\boldsymbol{M}[(\mathrm{Sr,Br,3} \rightarrow \mathrm{Ca,Br,3}),\Delta H_a]$ are almost as close. The one farthest from being a unit matrix is:

$$\boldsymbol{M}\{[\boldsymbol{V}(\mathrm{Mg,I,3,}\Delta H_a(\mathrm{kJ/mol})] \rightarrow \boldsymbol{V}[\mathrm{Br,F,3,}\Delta H_a(\mathrm{kJ/mol})]\} =$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 2.532\exp 0.580 \end{pmatrix} \tag{41}$$

Figures 5 and 6 give vectors $\boldsymbol{U}$ and $\boldsymbol{V}$ for group-3 halides; this set includes the trihalides and therefore requires the use of Eqs. (10) to (12). Data for gallium and indium bromides and iodides are not available in Ref. 24. Figure 6 gives the vector indices $\boldsymbol{V}$ and Figure 7 shows the periodicities in $\boldsymbol{V}_2$.

Central atom	F	Cl	Br	I
B	$\begin{pmatrix} 0 \\ -742\pm 20 \\ -1217\pm 20 \\ -1925\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -570\pm 20 \\ -861\pm 20 \\ -1322\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -443\pm 20 \\ -695\pm 20 \\ -1096\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -353\pm 20 \\ -545\pm 20 \\ -856\pm 10 \end{pmatrix}$
Al	$\begin{pmatrix} 0 \\ -670\pm 10 \\ -1152\pm 60 \\ -1765\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -498\pm 10 \\ -857\pm 60 \\ -1269\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -432\pm 10 \\ -723\pm 60 \\ -1067\pm 10 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -366\pm 3 \\ -592\pm 60 \\ -835\pm 10 \end{pmatrix}$
Ga	$\begin{pmatrix} 0 \\ -580\pm 10 \\ -960\pm 60 \\ -1421\pm 20 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -461\pm 10 \\ -750\pm 60 \\ -1060\pm 10 \end{pmatrix}$		
In	$\begin{pmatrix} 0 \\ -510\pm 10 \\ -870\pm 60 \\ -1327\pm 60 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -431\pm 10 \\ -700\pm 60 \\ -975\pm 10 \end{pmatrix}$		

Ligand atom

Figure 5. Archive vectors for ΔH_a data of boron-group halogens.

Central atom	F	Cl	Br	I
B	$\begin{pmatrix} 0 \\ 0.639\exp 8.92 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.447\exp 13.00 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.366\exp 10.38 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.286\exp 11.58 \end{pmatrix}$
Al	$\begin{pmatrix} 0 \\ 0.591\exp 6.60 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.430\exp 7.21 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.363\exp 8.54 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.290\exp 14.14 \end{pmatrix}$
Ga	$\begin{pmatrix} 0 \\ 0.483\exp 9.11 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.367\exp 13.62 \end{pmatrix}$		
In	$\begin{pmatrix} 0 \\ 0.465\exp 9.25 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0.340\exp 15.30 \end{pmatrix}$		

Ligand atom

Figure 6. Vector indices for ΔH_a data of boron-group halogens.

Graphs of raw data for transition-metal oxides are shown in the lower portion of Figure 4 of Ref. 7. Some of the original data are unavailable,[26] so the points have been digitized and the vector indices can be made available on request. They increase monotonically for group-4 to group-8 central atoms of periods 4 to 6. The curvatures are slightly positive as the number of oxygens increases, due to ligand-ligand repulsion.[7]

At this point, we demonstrate that another molecular property, ΔH_f° (298.15 K), can be treated using the methods from this paper. Uranium atoms exist in the crystalline

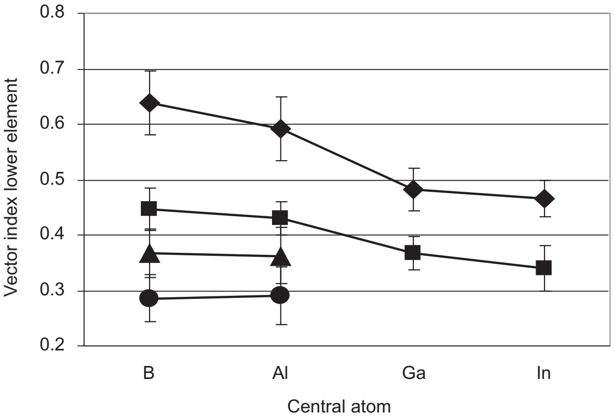

Figure 7. Same as Figure 4, except for showing the variation of $\boldsymbol{V}_2$ magnitudes for ΔH_a data of halides of group-3 atoms.

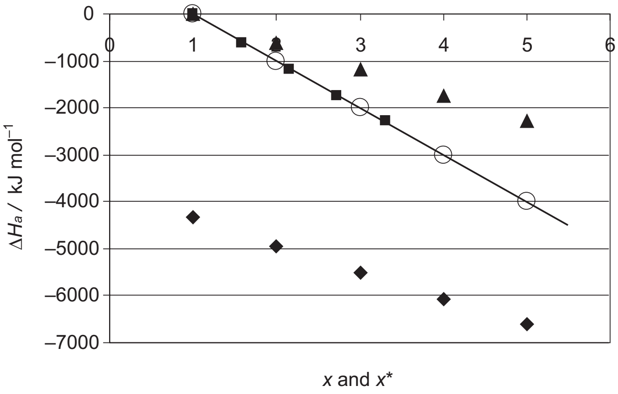

Figure 9. Same as Figure 8, except for P_4O_6 to P_4O_{10}. ▲: normalized data.

phase at STP, so their ΔH_f° (298.15 K)[27] is zero and the series belongs to Case 1. This property of crystalline uranium oxides is plotted in Figure 8. It is clear that there is serious departure from linearity. It would be unwise to define a vector index and even worse to *forecast* the missing datum for UO by using the otherwise suitable Eqs. (34) and (35).

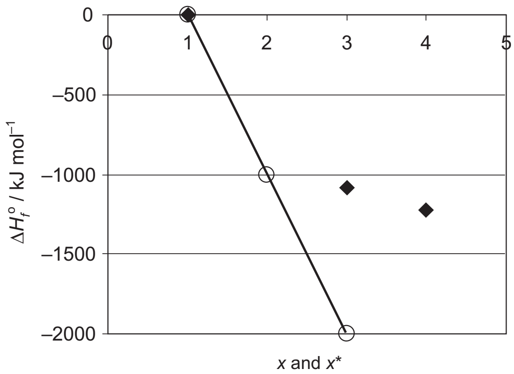

Figure 8. ◆: Tabulated ΔH_f° (298.15 K) data for solid uranium and its oxides plotted against x, the total number of atoms in the molecules. O: Data and line for the standard atom A and standard molecules AL and AL_2.

*Case 2: X at (1,*y*) and Subsequent Points with Negative Ordinates*

From here onward we no longer call upon related sets of molecular series but take up isolated series as needed to introduce variations of the fundamental theory and to illustrate its usefulness for different properties.

First we introduce a variation of the theory: the central »atom« A is allowed to become a molecule, so even if this »atom« is in the gas phase, its ΔH_a is not zero and any such series belongs to Case 2. Take the molecule P_4O_6, for instance (Figure 9). It has an adamantane structure with a phosphorus atom at each »vertex« and an oxygen atom in each »edge«.[28] From one to four oxygen atoms may be bonded to its vertices, giving the series P_4O_6 to P_4O_{10}.

$$\boldsymbol{U}(\mathrm{P_4O_6,O,5,}\Delta H_a) = \begin{pmatrix} -4326.26 \pm 60 \\ -4948.7 \pm 20 \\ -5510.5 \pm 20 \\ -6063.3 \pm 20 \\ -6600.1 \pm 6 \end{pmatrix} \quad (42)$$

The vector index is

$$\boldsymbol{V}(\mathrm{P_4O_6,O,5,}\Delta H_a) = \begin{pmatrix} -4326.26 \pm 20 \\ 0.292 \exp 5.06 \end{pmatrix} \quad (43)$$

A similar instance is N_2O_3. It has the form ONONO.[28] One and then two additional oxygen atoms can be bonded to the nitrogen atoms to form a series of three species[24] ending in N_2O_5. The archive vector is:

$$\boldsymbol{U}(\mathrm{N_2O_3,O,3,}\Delta H_a) = \begin{pmatrix} -1590.79 \pm 0.6 \\ -1908.37 \pm 2 \\ -2152.67 \pm 2 \end{pmatrix} \quad (44)$$

These data are shown in Figure 10, which resembles Figure 9. The fitted points x^* show less scatter than the original points x; this situation will prevail except when the *normalized* data lie to the left of the trend-line for the standard atom and ligands.

The vector index for the data in Eq. (44) is:

$$\boldsymbol{V}(\mathrm{N_2O_3,O,3,}\Delta H_a) = \begin{pmatrix} -1590.79 \pm 0.6 \\ 0.288 \exp 7.29 \end{pmatrix} \quad (45)$$

We return to ΔH_f° (298.15 K) again. Data for this property of gaseous group-2 halides are given in the JANAF tables.[29] The molecules are in the gaseous

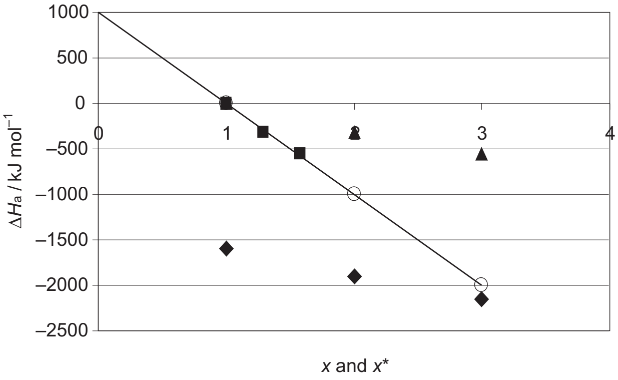

Figure 10. Same as Figure 8, except for N_2O_3 to N_2O_5.

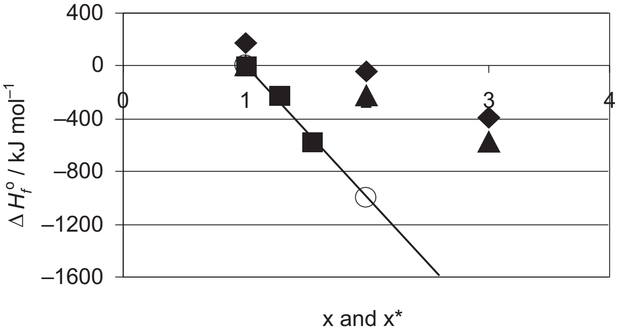

Figure 11. Same as Figure 8, except for ΔH_f^o (298.15 K) data of calcium chlorides.

Central atom	F	Cl	Br	I
Be	$\begin{pmatrix}320\pm5\\ -169.9\pm8\\ -796.01\pm4.2\end{pmatrix}$	$\begin{pmatrix}320\pm5\\ 60.7\pm12.6\\ -360.2\pm10.5\end{pmatrix}$	$\begin{pmatrix}320\pm5\\ 120.1\pm41.8\\ -229.3\pm16.7\end{pmatrix}$	$\begin{pmatrix}320\pm5\\ 170\pm41.8\\ -64\pm33.5\end{pmatrix}$
Mg	$\begin{pmatrix}147.1\pm0.8\\ -236.81\pm8.4\\ -726.76\pm3.1\end{pmatrix}$	$\begin{pmatrix}147.1\pm0.8\\ -43.5\pm42\\ -392.46\pm2.1\end{pmatrix}$	$\begin{pmatrix}147.1\pm0.8\\ -35.3\pm41.8\\ -302.9\pm10.5\end{pmatrix}$	$\begin{pmatrix}147.1\pm0.8\\ 24.61\pm41.8\\ -160.25\pm10.5\end{pmatrix}$
Ca	$\begin{pmatrix}177.8\pm0.8\\ -272\pm8\\ -784.5\pm8\end{pmatrix}$	$\begin{pmatrix}177.8\pm0.8\\ -104.6\pm13\\ -471.5\pm4\end{pmatrix}$	$\begin{pmatrix}177.8\pm0.8\\ -49.4\pm41.8\\ -384.9\pm8.4\end{pmatrix}$	$\begin{pmatrix}177.8\pm0.8\\ -5.04\pm84\\ -258.2\pm17\end{pmatrix}$
Sr	$\begin{pmatrix}164\pm1.7\\ -294.54\pm8.4\\ -766.09\pm4.2\end{pmatrix}$	$\begin{pmatrix}164\pm1.7\\ -123.85\pm8.4\\ -473.21\pm6.3\end{pmatrix}$	$\begin{pmatrix}164\pm1.7\\ -89.1\pm42\\ -407.1\pm12.6\end{pmatrix}$	$\begin{pmatrix}164\pm1.7\\ -30.49\pm83.7\\ -274.89\pm6.3\end{pmatrix}$

Ligand atom

Figure 12. Archive vectors for ΔH_f^o (298.15 K) data of boron-group halogens.

Central atom	F	Cl	Br	I
Be	$\begin{pmatrix}320\\ 0.544\exp 7.40\end{pmatrix}$	$\begin{pmatrix}320\\ 0.324\exp 14.81\end{pmatrix}$	$\begin{pmatrix}320\\ 0.26\exp 17.31\end{pmatrix}$	$\begin{pmatrix}320\\ 0.184\exp 13.59\end{pmatrix}$
Mg	$\begin{pmatrix}147.1\\ 0.426\exp 7.44\end{pmatrix}$	$\begin{pmatrix}147.1\\ 0.254\exp 18.90\end{pmatrix}$	$\begin{pmatrix}147.1\\ 0.216\exp 12.04\end{pmatrix}$	$\begin{pmatrix}147.1\\ 0.147\exp 12.93\end{pmatrix}$
Ca	$\begin{pmatrix}177.8\\ 0.468\exp 8.53\end{pmatrix}$	$\begin{pmatrix}177.8\\ 0.272\exp 14.06\end{pmatrix}$	$\begin{pmatrix}177.8\\ 0.235\exp 6.98\end{pmatrix}$	$\begin{pmatrix}177.8\\ 0.166\exp 5.73\end{pmatrix}$
Sr	$\begin{pmatrix}164\\ 0.464\exp 0.86\end{pmatrix}$	$\begin{pmatrix}164\\ 0.312\exp 5.77\end{pmatrix}$	$\begin{pmatrix}164\\ 0.279\exp 6.81\end{pmatrix}$	$\begin{pmatrix}164\\ 0.214\exp 7.01\end{pmatrix}$

Ligand atom

Figure 13. Vector indices for ΔH_f^o (298.15 K) data of boron-group halogens.

phase, just as for ΔH_a, but now the reference phases of the central and ligand species are metal atoms and gaseous dimers, so the gas phase atom datum is not zero. Figure 11 shows a representative plot of points for the original, normalized, and fitted data for calcium halides. Figures 12 and 13 show the vector sets $\boldsymbol{U}$ and $\boldsymbol{V}$. Figure 14 shows the trends visible among $\boldsymbol{V}_2$ of the vector indices; the enthalpies of formation show periodicity in the vertical ordering of the curves and secondary periodicity in their vertical spacings. The ΔH_f^o (298.15 K) have less clear manifestation of periodicity than ΔH_a.

We introduce another variation of the fundamental theory, the *substitution* of ligands in a central molecule. The example will be liquid chlorobenzenes; the data for them are described by:

$$\boldsymbol{U}[\mathrm{C_6H_6,H{\rightarrow}Cl},4,\Delta H_f^o(298.15\ \mathrm{K})] = \begin{pmatrix}49.1\\ 11.1\\ -19.1\pm1.6\\ -63.1\end{pmatrix} \tag{46}$$

where H→Cl indicates the substitution, $\boldsymbol{V}_3$ is the average of two of the three possible values for dichlorobenzene, and $\boldsymbol{V}_4$ is just one of the three possible values for trichlo-

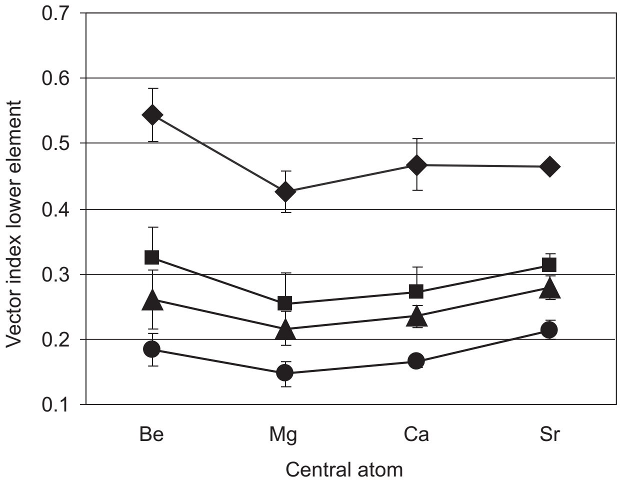

Figure 14. Same as Figure 4, except for showing the variation of $\boldsymbol{V}_2$ magnitudes in Figure 13 for ΔH_f^o (298.15 K) data of halides of group-3 atoms.

robenzene. The graph of these data resembles Figure 11. Using Eqs. (13) to (16) we find:

$$\boldsymbol{V}[\mathrm{C_6H_6,H{\rightarrow}Cl,4,}\Delta H_f^{\mathrm{o}}(298.15\ \mathrm{K})] = \begin{pmatrix} 49.1 \\ 0.368 \exp 6.30 \end{pmatrix} \tag{47}$$

Heats of combustion[28] for nitroalkanes are the final example of Case 2. The vector index is:

$$\boldsymbol{V}[\mathrm{CH_3N_2,H_2,5,}\Delta H_c^{\mathrm{o}}(\mathrm{kJ/mol})] = \begin{pmatrix} -709.6 \\ 0.653 \exp 0.358 \end{pmatrix} \tag{48}$$

where the very small errors of the tabulated data (0.056 % maximum) are ignored.

Case 4: X at (1,y) and Subsequent Points with Positive Ordinates

No data have been found and analyzed for Case 3, so we turn to Case 4. Uranium compounds are again employed to introduce a new property. The total entropy S^{o}_{298} of crystalline uranium oxides are plotted in Figure 15, as tabulated,[27] and as normalized and reflected through the x axis. Figure 16 shows the data fitted to line A to AL_2. The greatly expanded horizontal scale brings out a mild non-linearity of the fitted data that is not at all visible in Figure 15.

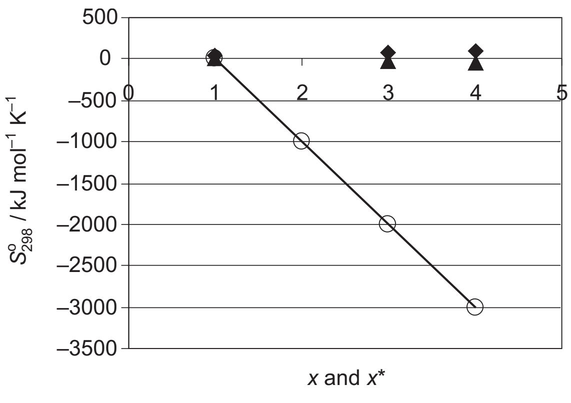

Figure 15. Same as Figure 8, except for tabulated, and normalized and reflected data for S° (298 K) of solid U, UO_2, and UO_3.

The archive vector and vector indices for S^{o} (298 K) of solid U, UO_2, and UO_3.are:

$$\boldsymbol{U}[\mathrm{U,O,4,}S^{\mathrm{o}}(298.15\ \mathrm{K})] = \begin{pmatrix} 50.2 \\ \mathrm{N.A.} \\ 77 \\ 96.1 \end{pmatrix} \tag{49}$$

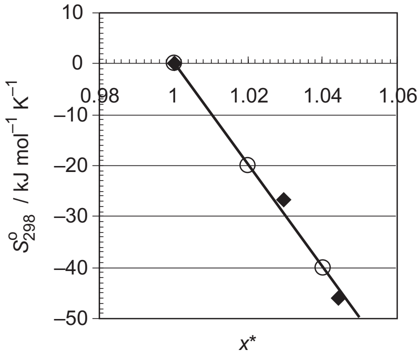

Figure 16. Same as Figure 15, except that the normalized and reflected data for S° (298 K) of solid U, UO_2, and UO_3 have been fitted in the least-squares fashion to the line of A, AL, and AL_2 and that both the x and y axes have been rescaled.

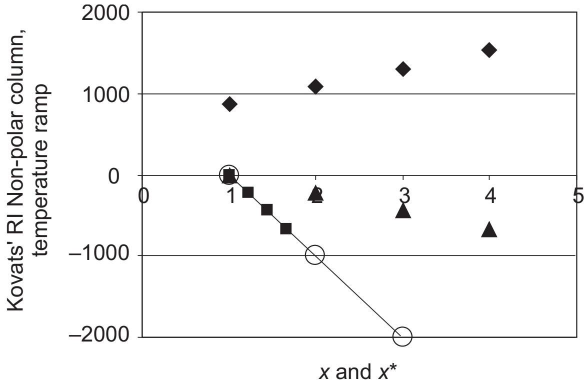

Figure 17. Kovats' gas-chromatography retention index (RI) for a non-polar column (30m/0.25mm/0.25 μm, N_2) with temperature ramp (40C@3min, 4K/min 240C@10min) measurements on diallyl monosulfide through tetrasulfide.

$$\boldsymbol{V}[\mathrm{U,O,4,}S^{\mathrm{o}}(298.15\ \mathrm{K})] = \begin{pmatrix} 50.2 \\ 0.1047 * \exp 7.45 \end{pmatrix} \tag{50}$$

Kovats' retention index (RI) values for gas-phase diallyl mono, di, tri, and tetrasulfides, as measured by Kubec *et al.*[30] are shown in Figure 17. The vectors are:

$$\boldsymbol{U}[\mathrm{C_6H_{10}S,S,4,RI(30m/0.25mm/0.25\mu m,N_2,40C@3min,4K/min,240C@10min)}] = \begin{pmatrix} 861 \\ 1079 \\ 1297 \\ 1538 \end{pmatrix} \tag{51}$$

$$\boldsymbol{V}[\mathrm{C_6H_{10}S,S,4,RI(30m/0.25mm/0.25\mu m,N_2,40C@3min,4K/min,240C@10min)}] = \begin{pmatrix} 861 \\ 0.223 \exp 3.44 \end{pmatrix} \tag{52}$$

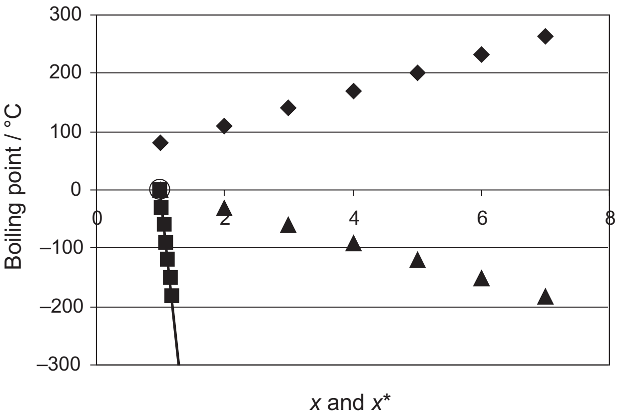

Figure 18. Same as Figure 8, but for boiling points of benzene with zero to six methyl substitutions.

Ray and colleagues[31] studied the hydrophobicity (log P) of the barbiturates.

$$\boldsymbol{U}(\mathrm{C_{10}H_{14}N_2O_3},14\rightarrow16\rightarrow18,3,\log P) = \begin{pmatrix} 1.15 \\ 1.65 \\ 2.15 \end{pmatrix} \quad (53)$$

The data, if plotted, resemble Figure 15 except that they are *precisely linear.* Hence, the error in $\boldsymbol{V}_2$ is exactly zero:

$$\boldsymbol{V}(\mathrm{C_{10}H_{14}N_2O_3},14\rightarrow16\rightarrow18,3,\log P) = \begin{pmatrix} 1.15 \\ 0.0005 * \exp 0.000 \end{pmatrix} \quad (54)$$

The well-known boiling points of methylbenzenes[32] are seen rather dramatically in Figure 18 and the vector index is:

$$\boldsymbol{V}(\mathrm{C_6H_6,H\rightarrow CH_3},7,\mathrm{BP}) = \begin{pmatrix} 80 \\ 0.303 * \exp 2.39 \end{pmatrix} \quad (55)$$

Generality of Matrix Transformation

We conclude by illustrating how powerful the matrix transformation process is. It can transform vector indices for very different series of molecules. The *species*, the *dimensionalities,* the *properties*, and the *phases* can be quite different. (The same holds for tabulated data except that the dimensionalities are restricted to the smaller one.) We show the matrix that transforms Eq. (48), for the heats of combustion of some liquid fuels, into Eq. (39), for the heats of atomization of gaseous phosphorus and two of its oxides. Determination of this matrix follows Eqs. (26) to (28).

$$\boldsymbol{M}\{[\boldsymbol{V}(\mathrm{CH_3N_2,H_2},5,\Delta H_c^{\mathrm{o}}(\mathrm{kJ/mol})]\rightarrow \boldsymbol{V}[\mathrm{N,O},3,\Delta H_a(\mathrm{kJ/mol})]\} = \begin{pmatrix} 1 & 1281 \exp 5.13 \\ 0 & 0.848 * \exp 5.49 \end{pmatrix} \quad (56)$$

DISCUSSION

In the Results, Case 2, missing data for NON and NONO were interpolated (to the left beyond the starting point of the series, N_2O_3) along the line AL_3 to A, after fitting known points to that line showed that they had very little scatter. In the same section, Case 4, there is a missing data point in the series U to UO_3, so Eqs. (34) and (35) must be used to find the difference between the abscissae for adjacent molecules. This difference, applied with the line A to AL_3, can be used to find the abscissa (x^{o}) and then the ordinate for UO; the result is a prediction of about (–15 + 50.2) kJ/mol K. Figure 19 shows that this forecast is not a good one. The abscissa difference, applied with the quadratic trend line through the points in Figure 19, gives a better value of (–12 + 50.2) kJ/mol K. (Caveat: the absence of data for solid UO suggests that it may not exist under normal conditions.) Other forecasts for data obtained in this way are given in Ref. 33. It might be possible to define a quadratic vector index (with three elements) for these and other seriously non-linear data and to compare them for evidences of periodicity.

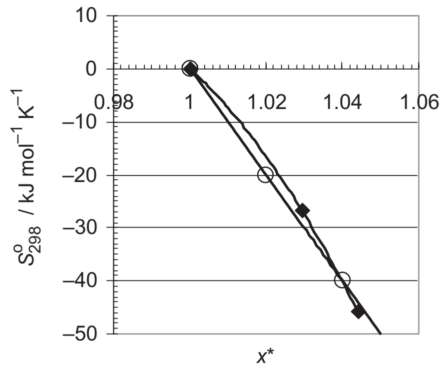

Figure 19. Same as Figure 16 except that a quadratic fit to the data has been used to make a very approximate forecast for UO.

It is interesting that the study of molecular similarity, while now using very sophisticated methods,[34] once had a simple approach somewhat parallel to the one in this paper. Karapet'yantz[35] found many closely linear relationships between data values for series of molecules plotted against data values for other series of atoms or molecules. Examples are the standard entropy of group-14 tetrachlorides *vs.* the standard entropy of group-14 tetrabromides, and the enthalpy of the following reaction for group-2 metals E(II)

$$E(II)(OH)_2\ (s) + CO_2\ (g) \rightarrow E(II)(CO_3)\ (s) + H_2O\ (l) \quad (57)$$

vs. the enthalpy of a similar reaction for group-1 metals E(I)

$$E(I)_2OH\ (s) + CO_2\ (g) \rightarrow E(I)_2CO_3\ (s) + H_2O\ (l) \quad (58)$$

Many of his plots show the effects of periodicity.

NOTE ADDED IN PROOF

Dr. Henry Kuhlman has pointed out that there are alternatives to the inverse least squares derivation given in this paper. One is to plot the data on the x axis and the number of molecular atoms on the y axis, and then to use suitable software to do a least squares fit with the y intercept locked at 1.0. The other is to use the number of added or substituted species n as a temporary independent variable and to do a software fit that passes through the origin.

Acknowledgement. – The authors thank Rebecca Chung for her valuable assistance.

REFERENCES

1. S. A. Shchukarev, *Vestn. Leningr. Univ.* **28** (1957) 795–811.
2. S. A. Shchukarev, *Zh. Obshchei Khim.*, **28** (1958) 845–859.
3. S. A. Shchukarev, *Neorganicheskaya khimiya*, Vol. 1, Vysshaya shkola, Moscow, 1970.
4. S. A. Shchukarev, *Neorganicheskaya khimiya*, Vol. 2, Vysshaya shkola, Moscow, 1974.
5. V. A. Latysheva, *Vodno-solevye rastvory. Systemnyi podkhod*, Saint Petersburg University, St. Petersburg, 1998.
6. V. A. Latysheva, *Zh. Obshchei Khim.*, **57** (1987) 2167–2179.
7. V. A. Latysheva and R. Hefferlin, *J. Chem. Inf. Comput. Sci.* **44** (2004) 1202–1209.
8. H. Bögel, J. Dettmann, and M. Randić, *Croat. Chem. Acta* **70** (1997) 827–840.
9. R. J. Quann and S. B. Jaffe, *Ind. Eng. Chem. Res.* **31** (1992) 2483–2497.
10. F. A. Kong, *J. Mol. Struct. (THEOCHEM)* **90** (1982) 17–28.
11. R. A. Hefferlin, *J. Chem. Inf. Comput. Sci.* **34** (1994) 314–317.
12. F. A. Kong, in: R. Hefferlin (Ed.), *Periodic Systems and their Relation to the Systematic Analysis of Molecular Data,* Edwin Mellen Press, Lewiston, New York, 1989, Chapter 11. pp. 396–413.
13. C. M. Carlson, R. J. Cavanaugh, R. A. Hefferlin, and G.V. Zhuvikin, *J. Chem. Inf. Comp. Sci.* **36** (1996) 396–398.
14. R. A. Hefferlin, G. V. Zhuvikin, K. E. Caviness, and P. J. Duerksen, *J. Quant. Spectrosc. Radiat. Transfer* **32 (**1984) 257–268.
15. F. A. Kong, in preparation.
16. E. Scerri, *Synthese,* **111** (1997) 213–232.
17. R. Scerri, *Found. Chem*. **6** (2004) 93–116.
18. M. Randić, *J. Chem. Ed.* **69** (1993) 713–718.
19. M. Razinger, J. R. Chrétien, and J. E. Dubois, *J. Chem. Inf. Comput. Sci.* **25** (1985) 23–27.
20. S. C. Basak, G. D. Grunwald, B. D. Gute, K. Balasubramanian, and D. Opitz, *J. Chem.Inf. Comput. Sci*. **40** (2000) 885–890.
21. S. Nikolić, G. Kovačević, A. Miličević, and N. Trinajstić, *Croat. Chem. Acta* **76** (2003) 113–124.
22. T. Balaban, *Rev. Roum. Chim.* **18** (1973) 635–653.
23. T. Balaban, *MATCH Commun. Math. Comput. Chem.* **21** (1986) 115–122.
24. L. V. Gurvich (Ed.), *Termodinamicheskie svoistva individual'nykh veschestv,* Vols. 1A-4B, Nauka, Moscow, 1978, 1979, 1981, and 1982.
25. S. Nikolić, A. Miličević, and N. Trinajstić, *Croat. Chem. Acta* **78** (2005) 241–250.
26. V. A. Latysheva, St. Petersburg University, Russia, private communication, September 2004.
27. D. R. Lide (Ed.), *CRC Handbook of Chemistry and Physics*, 84th edition, CRC Press, Boca Raton, Florida, 2003.
28. http://webbook.nist.gov/chemistry/
29. M. W. Chase, C. A. Davies, J. R. Downey, D. J. Frurip, R. A. McDonald, and A. N. Syverud, *JANAF Thermochemical Tables*, Third Edition, Parts I and II. *J. Phys. Chem. Ref. Data*, **14** (1985) Supplement 1.
30. R. Kubec, J. Velísek, M. Dolezal, and V. Kubelka, *J. Agric. Food Chem.* **45** (1997) 3580–3585.
31. S. K. Ray, S. C. Basak, C. R. Chandhury, A. B. Ray, and J. J. Ghosh, *Acta Ciencia Indica* **7** (1981) 187–192.
32. D. J. Klein and L. Bytautas, *MATCH Commun. Math. Comput. Chem.* **42** (2000) 261–290.
33. R. Hefferlin, *Proceedings of the 9th WSEAS International CSCC Multiconference,Vouliagmeni, Athens, July 11-16,* ISBN 960-8457-29-7 compact disk edition, 2004.
34. M. Johnson and G. Maggiora (Eds.), *Concepts and Applications of Molecular Similarity*, Wiley-Interscience, New York, 1990.
35. M. Kh. Karapetýantz, *Khimicheskoe srodstvo*, Znanie, Moscow, 1972.

SAŽETAK

Vektorski indeks osjetljiv na periodičnost za male molekule

Ray Hefferlin i Ken Luk

U beskrajnom molekularnom prostoru postoje mnoge serije molekula tipa X, XY_n, …, XY_N, gdje je N neki broj između broja 3 i nekoga broja nešto manjega od 10. Podatci za svojstva i faze tih molekula su aproksimativno lienarni obzirom na n. Autori su razvili vektorski prikaz tabuliranih podataka u seriji i vektorski indeks za opis serije. Započeli su s X kao metalnim atomom i sa svojstvom toplina atomizacije i pokazali su da vektorski indeks iskazuje periodičnost. Nakon toga su razmatrali slučajeve kada je X molekula, a svojstva su entalpija nastajanja, entropija, indeks retencije, hidrofobičnost i vrelište. Vektorski indeks je dvo-dimenzionalni vektor, čiji gornji element prikazuje svojstvo atoma ili molekule X, a donji element prikazuje razliku između bilo koja dva člana serije nakon što su podatci podešeni pomoću metode najmanjih kvadrata prema standardnoj, i linearnoj obzirom na n, seriji A, AL, …, AL_n, AL_N. Matrice mogu transformirati vektore podataka bilo koje serije u bilo koju drugu seriju istih dimenzija. Matrice mogu također transformirati vektorski indeks za bilo koje podatke o svojstvima neke aproksimativno linearne serije u bilo kojoj fazi u vektorski indeks za bilo koju drugu aproksimativno linearnu seriju.