Statistical Aspects of Second and Third Law Heats

William S. Horton

Institute for Materials Research, National Bureau of Standards, Washington, D.C. 20234

The effects of random errors and of nonconstant heat of vaporization upon the estimation of the second and third law heats of vaporization are examined. The most important conclusion is that the often noticed marked improvement in precision of the third law heat over that for the second law heat is real and is a natural consequence of the difference between the two estimators. The effects of systematic errors upon the two heats are not investigated here. Other results of interest, but of less importance because of the small magnitude of the effects, include: (1) the two heats are negatively correlated, (2) the second law heat is generally biased, (3) the third law heat is not the minimum variance unbiased estimator of the heat of vaporization, and (4) the standard deviation obtained from least squares fitting consistently overestimates the true standard deviation, but by a negligible amount. Any reversible process for which the equilibrium constant is treated by a similar procedure is governed by the same considerations. The results apply approximately at any temperature although the stimulus for these considerations comes from high temperature chemistry.

Key Words: Bias of least squares estimators, coefficient of variation, correlated estimators, enthalpy adjustments, heat of vaporization, high temperature chemistry, overestimation of standard deviation, second law heat, slope estimators, third law heat, vapor pressure.

1. Introduction

For some years it has been the practice when reporting high temperature vapor pressure data to compute and discuss "second law" and "third law" heats of vaporization or sublimation. Because each experimental temperature yields a third law heat, the individual values are often examined for consistency. Since each of these should be estimating the same quantity, grossly different values of one or of a few individuals are suggestive of experimental difficulty. With outliers thus considered, and generally discarded from future consideration, the remaining data are examined for trend, either with temperature or with chronological sequence. Trend with temperature is generally thought to be caused by systematic error in temperature measurement or in vapor pressure measurement, or by systematic error in the tabular free energy functions used to compute the third law heats. Trend with experimental sequence occasionally means that window or similar corrections used for optical pyrometry, were changing from run to run. This could be caused by darkening of the observation window as experimentation proceeds, for example. When the third law heats appear to pass this critical evaluation, the average is generally compared to the second law heat. The latter is usually obtained from the slope of a least-squares fit of the logarithm of the pressure to a straight line in the reciprocal of the absolute temperature. This comparison is also supposed to serve as a test of the experimentation [1] [1] ". . . because if the

heats calculated by the two procedures do not agree within the carefully calculated errors, then the experiment contains inconsistencies which all too often indicate either serious systematic errors or perhaps unknown species."

The uses described above to which the two estimates of the heat of transformation are put imply that those estimates are not biased by the method of estimation and furthermore that they are independent in the probability sense. It is the purpose of this article to point out the failure of these implicit conditions to apply and to examine the consequences of this failure. In general, the questions of unbiasedness and independence are always pertinent when two experimental values subject to random error are compared. Should either or both be biased by the means of computation, the difference may be due to that rather than to factors under study. Independence is always questionable when, as in the case of the second and third law heats, two estimates are computed using the same data. Lack of independence in the statistical sense would mean that random errors causing one estimate to be high would cause the other to be high also, in the case of positive correlation. Negative correlation would mean that random errors would have opposite effects on the two estimates. Independence would mean that there is no definite relation of this kind. The presence of correlation needs to be taken into account when a comparison is made.

It is found that the bias in the second law heat, as usually calculated, is generally negligible because of the currently attainable precision of most methods for measuring high temperature vapor pressures. The

---

[1] Figures in brackets indicate the literature references at the end of this paper.

lack of independence of the two heats does not significantly affect the size of the random error to be attributed to their difference. An important result is the understanding that the greater imprecision of the second law heat compared to that of the mean third law heat is a natural consequence of the difference between the two estimators and of the experimental temperatures usually involved. The usual method of computing the variance for the second law heat overestimates the variance when the simplest linear equation is fitted by least squares. This overestimation is quite small, however, and is naturally on the side of conservatism. Although the usually computed third law heat is not the minimum variance unbiased estimator, the loss of precision is quite small. The vaporization process is the main consideration here, but it is obvious that any reversible process for which the equilibrium constant is treated by a similar procedure is governed by the same considerations. Similar problems arise when estimating the entropy of transformation as the least-squares intercept of the usual second law treatment. As these problems will be obvious after considering the second law heat, entropy problems will not be treated here.

In what follows, the second and third law heats are discussed with a minimum of mathematical statistics. Statistical derivations, which are either already available elsewhere or are straightforward, are relegated to an appendix for statistically minded readers. Following the theoretical discussion of second and third law heats, a numerical example is given to illustrate the magnitudes of the effects.

2. Theoretical

2.1. Second Law Heat

Empirically, the vapor pressure, p, generally fits the equation

$$R \ln p = A - \frac{B}{T} \qquad (1)$$

within experimental error, where R is the gas constant, A and B are assumed to be constant and T is the absolute temperature. However, theoretically at a single temperature

$$R \ln p = \Delta S^\circ - \frac{\Delta H^\circ}{T} \qquad (2)$$

where ΔS° and ΔH°, the entropy and enthalpy changes, are slowly varying functions of T. The superscript $^\circ$ specifies the substances to be in standard states. Equation (2) assumes that the pressure is sufficiently low for the gas to be assumed ideal. One method of treating the data fits experimental values of $R \ln p$ to $1/T$ and adjusts $\hat{B}$, the least squares estimate of B in eq (1), to a reference temperature T_r. $\hat{B}$ might be considered a reasonable estimate of ΔH° for some representative temperature within the experimental

range; the enthalpy change at a commonly used reference temperature far removed from the region of the experiment is somewhat different. Adjustment of $\hat{B}$ to such a temperature is given by

$$\hat{b}_2 = \hat{B} - \int_{T_r}^{T_e} \Delta C_p^\circ dT \qquad (3)$$

where $\hat{b}_2$ is used to represent the estimate of the second law heat obtained by adjusting $\hat{B}$ from the experimental temperature, T_e, to the reference temperature, T_r. The quantity ΔC_p° is the heat capacity of the vapor minus the heat capacity of the condensed phase at constant pressure. ΔC_p° is a function of T, but is independent of p.

The reason, of course, that eq (1) is usable without serious error is that ΔC_p° is sufficiently small that its integral over a short range is negligible compared to ΔH°. The use of eq (3) is straightforward as long as one knows the correct value of T_e. Some workers make allowance for this problem by using heat capacity data as in the sigma method [2,3]. More often than not, however, the additional computational labor is not felt to be warranted and is not carried out. Instead, the experimenter chooses some temperature within the range of the experiment as corresponding to the least-squares slope in order that the adjustment by eq (3) may be carried out. Some authors use the arithmetic mean of the experimental temperatures. Others use the temperature corresponding to the arithmetic mean of the reciprocal temperatures. Still others use a rounded value somewhere near the midrange. However, it is shown below that if the data are required to fit eq (1), the quantity $\hat{B}$ does not represent unambiguously the heat of vaporization at any recognizable temperature within the range of experiment. Instead the quantity represents a combination of enthalpy and entropy terms and of the coefficients of the heat capacity equation.

It is particularly convenient in what follows to replace the variable $-T^{-1}$ by x. Then the exact vapor pressure equation may be written as

$$R \ln p = \alpha + \beta x + f(x), \qquad (4)$$

where a function of x (and therefore of T) has been added to the simplest form in order to make it exact. There is no other limitation on the form of this function. Then the usual least-squares slope for fitting eq (1) has as its expected value

$$E(\hat{B}) = \beta + \frac{\Sigma(x - \bar{x})f(x)}{\Sigma(x - \bar{x})^2} \qquad (5)$$

Alternatively it is possible to show that if $g(x)$ represents the slope of the secant connecting the point $(x, R \ln p)$ with the point on the curve at $\bar{x}$.

$$E(\hat{B}) = \frac{\Sigma(x-\bar{x})^2 g(x)}{\Sigma(x-\bar{x})^2}. \qquad (6)$$

Thus, the least squares linear slope is a weighted average of the true slopes of the set of secants where the weighting is proportional to the square of the distance from $\bar{x}$.

Because $\hat{B}$ does not represent the slope at any easily defined point, the usual choices for T_e in eq (3) lead to bias in $\hat{b}_2$, the second law heat. The bias, however, is quite small and is given by the integral of ΔC_p° from the chosen T_e to the correct, if unknown, value of T_e. Examples of its magnitude are given later. The main advantage of the modified procedures mentioned earlier appears to be that they avoid the ambiguity of the temperature to which the slope applies. Occasional extended discussions arise, however, about the "correct" temperature to use, and it is well to recognize that there is no easily defined one when the data are fitted directly to eq (1). The bias is not small if the lowest or highest experimental temperature is chosen rather than one near the midrange. Further discussion of this point is relegated to a later section. Removal of the small bias by some means may be considered desirable because, hopefully, vapor pressure data collection will improve over the years and, secondly, because comparison with third law heat for significant difference should be made with estimates known to be unbiased with respect to sources not being tested by the comparison.

Because the experimental data are fitted to a straight line and because theoretically the correct function is known to be more complex, the computed standard deviation is estimating a value larger than the true standard deviation. It is shown in the appendix that the increase in variance (the square of the standard deviation) is given by

$$EVF = (n-2)^{-1}\left\{\Sigma(f-\bar{f})^2 - \frac{[\Sigma(x-\bar{x})(f-\bar{f})]^2}{\Sigma(x-\bar{x})^2}\right\}, \quad (7)$$

where n is the number of observations and $f(x)$ has been shortened to f for convenience. This quantity, the extra variance function, is always positive and, as shown in the numerical example, is negligible for the type of experiments being considered.

2.2. Third Law Heat

Alternative to the second law treatment is that of the third law wherein for each experimental point, i, there is computed

$$b_{3i} = T_i(\delta_i - R \ln p_i), \qquad (8)$$

where $\delta_i = \Delta\left[-\dfrac{G_i^\circ - H_r^\circ}{T_i}\right]$, the change in free energy function for the process [4]. The free energy function, $-\dfrac{G_i^\circ - H_r^\circ}{T_i}$, for each substance is generally available in

tabulations or may be computed from other data. The b_{3i} are examined individually, as suggested in the introduction, and they are also averaged to give

$$\hat{b}_3 = \frac{1}{n}\Sigma b_{3i}, \qquad (9)$$

the estimator for the third law heat.

This estimator is unbiased. However, if the random errors are essentially constant when measured in $R \ln p$ at different temperatures, random errors in the b_{3i} will be a function of T because of the term $RT \ln p$. As a consequence $\hat{b}_3$ as given by eq (9) is not the estimate with the least variance. There will generally be a larger error associated with $\hat{b}_3$ than with another estimator, $\bar{b}_3$, where

$$\bar{b}_3 = \frac{\Sigma x(R \ln p - \delta)}{\Sigma x^2} = \frac{\Sigma T^{-1}(\delta - R \ln p)}{\Sigma T^{-2}}.$$

This is an unbiased estimator also. It can be shown that $\bar{b}_3$ is, of all unbiased estimators of the heat, the one with minimum variance and may be called the "minimum variance unbiased estimator" (MVUE). Because $\bar{b}_3$ has a smaller variance than $\hat{b}_3$, it might be preferred. However, it will be shown in the numerical example that for high temperature experiments the improvement in precision is small and does not warrant the loss of convenient computation. As already pointed out, the individual b_{3i} serve a useful purpose.

2.3. Similarity of Formulas for the Estimators

The estimators $\hat{b}_2$, $\hat{b}_3$, and $\bar{b}_3$, each of which is an estimate of ΔH_r°, can be written in similar notation and the similarity is striking when these are placed side by side as follows:

$$\left.\begin{array}{l} \hat{b}_2 = \dfrac{\Sigma(x-\bar{x})(R \ln p - \delta)}{\Sigma(x-\bar{x})^2} = \dfrac{\Sigma(T^{-1}-\overline{T^{-1}})(\delta - R \ln p)}{\Sigma(T^{-1}-\overline{T^{-1}})^2} \\[3mm] \hat{b}_3 = \dfrac{\Sigma x^{-1}(R \ln p - \delta)}{n} = \dfrac{\Sigma T(\delta - R \ln p)}{n} \\[3mm] \bar{b}_3 = \dfrac{\Sigma x(R \ln p - \delta)}{\Sigma x^2} = \dfrac{\Sigma T^{-1}(\delta - R \ln p)}{\Sigma T^{-2}}. \end{array}\right\} (10)$$

Although they are not treated in this article the effects of systematic errors, such as in the thermodynamic data, may be readily investigated using these equations. The comparison of the different estimators is particularly facilitated.

2.4. Variance and Covariance of the Estimates

Because the second and third law heats are obtained by an essentially statistical estimation procedure, the statistical properties of the estimators are of interest.

535

Of greatest interest are the expected values of the estimators, their variances, the covariance and the correlation coefficient. Knowledge of the law governing the random errors would be desirable, but is not necessary. At this time it can only be stated as likely that the random deviations from eq (1) are of the Gaussian type when measured as deviations in $R \ln p$. That errors are often unwarrantedly assumed to have this characteristic has been shown by Clancey [5]. For least-squares fitting it is merely necessary that the random errors have mean zero, constant variance and be independent of one another in the probability sense. An examination of representative data is proceeding in this laboratory to ascertain the validity of the normal assumption but sufficient information is not yet available. At this point, it does seem very likely that the magnitude of random errors in the logarithm of the pressure is not temperature dependent.

If the vapor pressure data are treated by the sigma method or an alternate, then the slope obtained by the second law is for an unambiguous temperature. Correspondingly, there will be no bias in the second law heat. From the definitions [6] of expected values, variances, covariance and correlation coefficient, ρ, the following relations are found.

$$E(\hat{b}_2) = \Delta H_T^\circ \tag{11}$$

$$E(\hat{b}_3) = \Delta H_T^\circ \tag{12}$$

$$\mathrm{Var}\ (\hat{b}_2) = \sigma^2 / \Sigma (T^{-1} - \overline{T^{-1}})^2 \tag{13}$$

$$\mathrm{Var}\ (\hat{b}_3) = \sigma^2 \Sigma T^2 / n^2 \tag{14}$$

$$\mathrm{Var}\ (\bar{b}_3) = \sigma^2 / \Sigma T^{-2} \tag{15}$$

$$\mathrm{Covar}\ (\hat{b}_2, \hat{b}_3) = \sigma^2 (1 - \overline{T}\ \overline{T^{-1}}) / \Sigma (T^{-1} - \overline{T^{-1}})^2 \tag{16}$$

$$\rho(\hat{b}_2, \hat{b}_3) = \frac{n(1 - \overline{T}\ \overline{T^{-1}})}{[\Sigma T^2 \Sigma (T^{-1} - \overline{T^{-1}})^2]^{1/2}}. \tag{17}$$

In these equations σ is the standard deviation of the random errors in $R \ln p$. It follows from Cauchy's inequality that $\overline{T}\ \overline{T^{-1}}$ is always greater than or equal to unity so that the second and third law heats are always negatively correlated in the statistical sense. In most high temperature experiments the degree of correlation is quite small as illustrated by an example later.

Because relative values are more informative, the following ratios are of interest:

$$\frac{\mathrm{Var}\ (\hat{b}_3)}{\mathrm{Var}\ (\hat{b}_2)} = \frac{\Sigma T^2 \Sigma (x - \bar{x})^2}{n^2} \tag{18}$$

$$\frac{\mathrm{Cov}\ (\hat{b}_2, \hat{b}_3)}{\mathrm{Var}\ (\hat{b}_2)} = 1 + \overline{T}\bar{x} \tag{19}$$

$$\frac{\mathrm{Var}\ (\hat{b}_3)}{\mathrm{Var}\ (\bar{b}_3)} = \frac{\Sigma T^2 \Sigma x^2}{n^2}. \tag{20}$$

(In these equations both T and x have been used for compactness.) All of these quantities have identical powers of T and x multiplying one another. Consequently, the magnitude of the mean absolute temperature of the experiments does not greatly affect the magnitude of the ratios. Instead these quantities and the correlation coefficient are functions of the coefficient of variation of the temperatures $CV(T)$, where this represents the relative spread among the temperatures chosen for the experiments rather than relative errors in the temperature. In fact, it can be shown that for moderately small values of $CV(T)$ the following are reasonable approximations.

$$\left. \begin{array}{c} \rho(\hat{b}_2, \hat{b}_3) \approx -CV(T) \\[2mm] \dfrac{\mathrm{Var}\ (\hat{b}_3)}{\mathrm{Var}\ (\hat{b}_2)} \approx [CV(T)]^2 \\[2mm] \dfrac{\mathrm{Cov}\ (\hat{b}_2, \hat{b}_3)}{\mathrm{Var}\ (\hat{b}_2)} \approx -[CV(T)]^2. \end{array} \right\} \tag{21}$$

Trial calculations have shown that the approximation for ρ is high by only 2 percent at $CV(T) = 0.15$. At this same value the approximation for the ratio $\mathrm{Var}\ (\hat{b}_3)/\mathrm{Var}\ (\hat{b}_2)$ is low by 13 percent, and the covariance ratio approximation is low by 4 percent. The ratio given by eq (18) can be greater than unity, but this is only true for experiments for which the $CV(T)$ is about 0.6 or greater. Application of Cauchy's inequality to eq (20) shows this ratio to be always greater than unity, as it must be if $\hat{b}_3$ is the minimum variance unbiased estimator.

3. Illustrative Calculations

In order to illustrate the magnitude of the effects discussed above, data for the vapor pressure of tungsten [7] will be used. Table 1 shows pertinent data from that study together with third law heats as usually computed. Table 2 compares results from various methods of computing the heat of sublimation from the data of table 1. The number of figures shown is not indicative of uncertainty. The choice was made merely in order to have at least two digits for the smallest difference between values.

TABLE 1. *Vapor pressures and heats of sublimation of tungsten* [a]

Temperature	Vapor pressure	$\Delta H (298)$
°K	atm × 10⁶	kcal mole⁻¹
2786	5.42	203.7
2773	5.21	202.9
2679	1.26	203.5
2574	0.250	203.7
2925	50.8	203.9
3054	98.3	204.6
2614	0.534	203.0
3068	197	202.7
2934	40.0	203.0
3183	542	204.0

[a] Data from reference 7.

TABLE 2. *Comparison of sublimation heats for tungsten*

Second Law, uncorrected (T_1)[a]	201442.5 cal/mole[b]
Second Law, uncorrected $(\bar{T})$	201619.6
Second Law, uncorrected (T_λ)	201885.1
Second Law, corrected	201626.5
Third Law, usual	203474.9
Third Law, MVUE	203457.5

[a] The temperature indicated in the parentheses is that to which it was assumed the slope corresponded: $T_1 =$ lowest experimental temperature, $T_\lambda =$ highest temperature, $\bar{T} =$ arithmetic mean temperature.

[b] The number of figures shown is not indicative of the uncertainty. The choice was made merely in order that the smallest difference among values would have at least two digits. For all, the reference temperature was 298 °K.

In order to give an unbiased second law treatment, logarithmic vapor pressures were corrected by an equivalent to the sigma method. The equation

$$R \ln p - \Delta(S^\circ - S_e^\circ) + \frac{\Delta(H^\circ - H_e^\circ)}{T} = \Delta S_e^\circ - \frac{\Delta H_e^\circ}{T} \quad (22)$$

has a form such that the entropy and enthalpy functions are easily evaluated by computer using the OMNITAB language [8] and the JANAF tables [9]. The differences between second law heats from corrected and uncorrected logarithmic vapor pressures is seen to be of the order of hundreds of calories depending upon the temperature assumed to correspond to the slope. In the corrected case the "experimental temperature" was chosen to be the arithmetic mean of those listed in table 1, 2857 °K. If the slope of the uncorrected second law treatment is assumed to correspond to a temperature near $\bar{T}$, the table shows that for the tungsten work the error is of the order of only 10 calories.

Table 2 also shows the usual third law result and the minimum-variance estimator. The difference of 17 calories per mol (0.008%) is also too small to be of importance.

TABLE 3. *Variance functions*

Variance, Second Law heat[a]	1.76×10^6
Variance, Third Law heat, usual[a]	8.20×10^3
Variance, Third Law heat, MVUE[a]	8.05×10^3
Covariance $(\hat{b}_2,\hat{b}_3)$[a]	-8.12×10^4
Correlation coefficient	-0.07
Extra variance function[a]	6.24×10^{-1}

[a] Variances are relative to that for $R \ln p$.

Table 3 is probably of more interest because the improved precision of the third law heat over that for second law heat is very evident. The square root of the first two entries gives the ratio of standard deviations, about 15. It should be emphasized that this is a natural outcome of the methods of estimation and is not a result of imprecision that can be overcome by better experimentation. Also evident in table 3 is the small decrease in variance for the minimum variance estimator compared to that for the usual third law heat. As mentioned earlier there would be little gained and some definite loss in using this estimator.

The correlation coefficient for $\hat{b}_2$ and $\hat{b}_3$ given in table 3 is seen to be a small negative quantity. Because of this, the correlation is of no great importance when comparing second and third law heats. In fact, because the third law heat is also much more precise than the second law heat, the random error in the comparison is contributed almost solely by the latter quantity.

The extra variance function is truly negligible, showing that the use of the simpler method of straight line fitting for these experiments has no important effect upon the estimate of error.

4. Appendix

The question of bias in the second law heat as illustrated by eq (5) is treated by assuming that if the "true" equation is given by (4), experimental values of $R \ln p_i$, represented by y_i, are given by

$$y_i = \alpha + \beta x_i + f(x_i) + \epsilon_i, \quad (23)$$

where the ϵ_i are assumed to behave like random errors. Equation (5) results from application of the well-known formula for the slope of a linear least-squares fit to this equation and taking expected values.

Consider the curve giving the true values, Y_i, as a function of the experimental values, x_i. Let $g(x)$ be the slope of the secant joining any point (x, Y) with the point $(\bar{x}, Y(\bar{x}))$. Then for a given i

$$\frac{Y - Y(\bar{x})}{x - \bar{x}} = \frac{Y - \bar{Y}}{x - \bar{x}} + \frac{\bar{Y} - Y(\bar{x})}{x - \bar{x}} = g(x). \quad (24)$$

Multiplying by $(x - \bar{x})^2$ and summing over i yields

$$\Sigma(x - \bar{x})(Y - \bar{Y}) = \Sigma(x - \bar{x})^2 g(x). \quad (25)$$

Replacing Y by $y - \epsilon$, dividing by $\Sigma(x - \bar{x})^2$ and taking expected values gives

$$\frac{\Sigma(x - \bar{x})(y - \bar{y})}{\Sigma(x - \bar{x})^2} = \frac{\Sigma(x - \bar{x})^2 g(x)}{\Sigma(x - \bar{x})^2}. \quad (26)$$

The left-hand member is the expression used to calculate the least squares slope of a linear fit and the right hand member is that for eq (6).

When a straight line is fitted to data which are known to be nonlinear it is expected that the computed variance will be biased upwards. Although the derivation is straightforward, the expression for the excess appears not to have been presented before. For this purpose let $y = \hat{\alpha} + \hat{\beta}x + f(x) + \epsilon$ represent experimental values and let $\hat{y} = \hat{\alpha} + \hat{\beta}x$ represent estimates obtained by a least-squares linear fit. The ϵ are random errors with zero mean and variance equal to σ^2. Also let $x - \bar{x} = u$ and $\Sigma(x - \bar{x})^2 = \Sigma u^2 = S$. Then for a given i

$$y - \hat{y} = (\beta - \hat{\beta})u + (f - \bar{f}) + (\epsilon - \bar{\epsilon})$$

where for simpler writing $f(x)$ has been replaced by f. But since

$$\beta - \hat{\beta} = -\frac{\Sigma uf}{S} - \frac{\Sigma u\epsilon}{S},$$

$$y_i - \hat{y} = \left[-\frac{u_i\Sigma uf}{S} + (f_i - \bar{f}) \right] + \left[-u_i \frac{\Sigma u\epsilon}{S} + (\epsilon_i - \bar{\epsilon}) \right],$$

where subscript i has now been indicated. The brackets segregate two terms, one with error terms and one without. Thus,

$$y_i - \hat{y} = A_i + e_i.$$

To determine what the usually computed variance is actually giving we need the expected value of the sum of squares for $y - \hat{y}$, or

$$E\{\Sigma(y - \hat{y})^2\} = \Sigma A_i^2 + E(\Sigma e_i^2).$$

The first term of the right-hand member is

$$\Sigma A_i^2 = \Sigma(f - \bar{f})^2 - \frac{(\Sigma uf)^2}{S}.$$

The second term becomes

$$E(\Sigma e_i^2) \doteq \frac{(\Sigma u^2 f)^2}{S^2}\sigma^2 - \frac{2}{S}E\{\Sigma u(\epsilon - \bar{\epsilon})\Sigma u\epsilon\} + E\{\Sigma(\epsilon - \bar{\epsilon})^2\}$$

where in the right-hand member subscripts have been dropped as they are no longer needed for clarity. This expression is evaluated easily as $(n-2)\sigma^2$. Because the computed variance is given by $(n-2)^{-1}\Sigma(y - \hat{y})^2$, the expected value follows as

$$E\{\text{Var}(\hat{y})\} = \sigma^2 + (n-2)^{-1}\left[\Sigma(f - \bar{f})^2 - \frac{(\Sigma uf)^2}{S}\right].$$

The second term of the right-hand member is equivalent to eq (7) and has been termed by the author the extra variance function. From its equivalent,

$$\frac{\underset{ij}{\Sigma\Sigma}[u_i(f - \bar{f})_j - u_j(f - \bar{f})_i]^2}{(n-2)S},$$

wherein the summations over i and j are limited by $i \neq j$ and $j > i$, the extra variance function is seen to be always positive.

The last equation of (10) results from dividing eq (8) by T_i and minimizing the sum of squared residuals with respect to b_3. Equations (13), (14), and (15) are the results of taking the variances of the linear forms, (10). Equation (13) also follows from the well-known results for a straight line. Equation (16) follows from the expected value of $(\hat{b}_2 - \Delta H_r^\circ)(\hat{b}_3 - \Delta H_r^\circ)$. Equation

(17) is by definition given as

$$\text{Covar}(\hat{b}_2, \hat{b}_3)/\sqrt{\text{Var}(\hat{b}_2) \cdot \text{Var}(\hat{b}_3)}.$$

In order to obtain the approximations of eqs (21), the Taylor's expansion for x around $\bar{T}$ is useful:

$$x \approx -\frac{1}{\bar{T}} + \frac{(T - \bar{T})}{\bar{T}^2} - \frac{(T - \bar{T})^2}{\bar{T}^3} + \dots$$

From this one obtains

$$\bar{x} \approx -\frac{1}{\bar{T}} - \frac{\Sigma(T - \bar{T})^2}{n\bar{T}^3}$$

and

$$\Sigma(x - \bar{x})^2 \approx \frac{\Sigma(T - \bar{T})^2}{\bar{T}^4}.$$

From the expression for $\bar{x}$, the relation

$$1 + \bar{T}\bar{x} \approx -\frac{\Sigma(T - \bar{T})^2}{n\bar{T}^2} = -[CV(T)]^2$$

is obtained, which by virtue of eq (19), gives the third approximation in (21).

In the exact expression $\Sigma T^2 = \Sigma(T - \bar{T})^2 + n\bar{T}^2$ the sum of the squared differences may be neglected when the $CV(T)$ is small, and one can use $\Sigma T^2 \approx n\bar{T}^2$. Using both T and $\bar{x}$ the square of (17) may be written as

$$\frac{n^2(1 + \bar{T}\bar{x})^2}{\Sigma T^2 \Sigma(x - \bar{x})^2} \approx \frac{\Sigma(T - \bar{T})^2}{n\bar{T}^2} = [CV(T)]^2.$$

We note that taking the square root of both sides now gives ambiguity with regard to sign, but it has already been pointed out that this must be negative. Thus, the first approximation of (21) is obtained. The second approximation follows readily from expressions already used here.

---

5. References

[1] R. J. Ackerman and R. J. Thorn, Vaporization of Oxides in Progress in Ceramic Science, Volume 1, p. 42, ed. by J. E. Burke. (Pergamon Press, N.Y., 1961).

[2] K. K. Kelley, Contributions to the Data on Theoretical Metallurgy. III. The Free Energies of Vaporization and Vapor Pressures of Inorganic Substances, Bulletin 383, U.S. Dept. of Interior, Bur. of Mines, 1935.

[3] G. N. Lewis and M. Randall, Thermodynamics, 2d ed., p. 175, revised by K. S. Pitzer and L. Brewer (McGraw-Hill Book Company, New York, 1961).

[4] S. Glasstone, Thermodynamics for Chemists, pp. 306–309 (D. Van Nostrand Co., New York, N.Y., (1947)).

[5] V. J. Clancey, Nature 159, 339–40 (1947).

[6] A. M. Mood and F. A. Graybill, Introduction to the Theory of Statistics, 2d ed., ch. 5 (McGraw-Hill Book Company, New York, N.Y., 1963).

[7] R. Szware, E. R. Plante, and J. J. Diamond, J. Res. NBS 69A (Phys. and Chem.) No. 5, 417 (1965).

[8] J. Hilsenrath, G. G. Ziegler, C. G. Messina, P. J. Walsh, and R. J. Herbold, Omnitab, A Computer Program for Statistical and Numerical Analysis, NBS Handbook 101 (1966).

[9] JANAF Thermochemical Tables, Dow Chemical Co., Midland, Michigan, 1961, Tungsten crystal and gas tables dated December 31, 1961.