The Limiting Gas Density Method for Molecular Weights

R. T. Birge and F. A. Jenkins

Articles you may be interested in

Molecular Weight Dependency of Rheological Characteristics of Linear Low Density Polyethylene
J. Rheol. **29**, 323 (1985); 10.1122/1.549793

Theory of nematic systems of semiflexible polymers. I. High molecular weight limit
J. Chem. Phys. **76**, 3295 (1982); 10.1063/1.443325

Method for Calculating Average Molecular Weights and Molecular Weight Distributions of Polymeric
Materials from Sedimentation Equilibrium Experiments
J. Chem. Phys. **32**, 1739 (1960); 10.1063/1.1731013

Molecular Weight Comparisons from Density and XRay Data. The Atomic Weights of Calcium and
Fluorine
J. Chem. Phys. **13**, 383 (1945); 10.1063/1.1723968

The Determination of MolecularWeight Distribution in High Polymers by Means of Solubility Limits
J. Appl. Phys. **16**, 419 (1945); 10.1063/1.1707609

The Limiting Gas Density Method for Molecular Weights

R. T. BIRGE AND F. A. JENKINS, *University of California, Berkeley*
(Received February 12, 1934)

In this, the first of a series of papers, we give a general discussion of the limiting gas density method for the determination of molecular weight. The various methods that have been used, or might well be used for the reduction of the data, are compared in a systematic way, and the most desirable methods for such work are pointed out. These methods are then applied to selected data given in the literature, and are compared with the actual published calculations and results. The paper closes with certain conclusions drawn from the preceding analysis.

INTRODUCTION

ONE of us[1] recently made a photometric measurement of the relative intensity of the $C^{12}C^{12}$ and $C^{12}C^{13}$ Swan bands. This result, combined with Aston's determination of the mass of C^{12} and with the known abundance of the oxygen isotopes, leads directly to an atomic weight of carbon of approximately 12.01, as contrasted with the international atomic weight 12.00.

In searching for the cause of this discrepancy, we have made a critical study of certain previous determinations of the atomic weight of carbon. Most of the recent work has been carried out by the "limiting gas density" method. We have found, in this study, that a number of different methods have been employed in the reduction of the data, and in many cases the logical basis for the method is not clear to us. We have, accordingly, made a certain number of recalculations, and have in some cases obtained results appreciably different from those published. It has not been possible to make a complete investigation of all previous data. In many cases the published information is too incomplete to allow a recalculation. It does seem desirable, however, to give at this time a general outline of the limiting gas density method, with illustrations of the various types of calculation that can be used and have been used in work in this field. From the calculations we have already made, it would appear that the best value of the atomic weight of

[1] F. A. Jenkins and L. S. Ornstein, Proc. Amsterdam Acad. **35**, 1212 (1932).

carbon is appreciably higher than 12.00 and that the discrepancy with the band spectrum value, 12.01, may therefore not be serious.

The results that we have obtained will be presented in a series of papers. In this, the first paper of the series, we discuss the general principles of the limiting gas density method. The second paper of the series will be devoted to the limiting gas density of oxygen, and the resulting value of R, the normal mole volume. In the third paper we shall discuss the atomic weight of carbon, as determined especially by the gas density method.

THE LIMITING GAS DENSITY METHOD

We define the density of a gas, at any pressure p, by

$$\rho_p = w/v, \tag{1}$$

where w is the mass of gas used, and v is the measured volume. Now let

$$L_p = \rho_p/p = w/pv. \tag{2}$$

Then if $p = 1$, $L_p = \rho_p$, the density at unit pressure. This is regularly denoted as merely L. Moreover, if we had a perfect gas ($pv = $ constant), L_p would be constant regardless of the value of p. Actually the value of L_p is found, in the case of all real gases, to vary with p. Hence L_p is the density the real gas *would have*, at unit pressure, *if* it acted like a perfect gas between the pressure at which the density ρ_p is actually measured and unit pressure. This is a useful concept, for at $p = 0$ any gas becomes asymptotically a perfect gas. Hence the limiting value $L_{\lim}$ of L_p at $p = 0$ is the density at unit pressure that the actual

gas would have *if* it were a perfect gas. This quantity L_{lim} is of course a function of temperature, but if all work is done at 0°C, then the product of L_{lim} and the normal mole volume R is, by definition, the molecular weight M, i.e.,

$$M = L_{\text{lim}} \cdot R. \qquad (3)$$

As previously discussed by one of us,[2] R is the normal mole volume in liter-mole^{-1} units, with $g_{45} = 980.616$ cm·sec.$^{-2}$, the units commonly employed in gas density work.

Eq. (3) is the fundamental equation for the determination of molecular weights by this method. The value of R is derived from the limiting density of oxygen, as will be discussed by one of us in the next article of this series.[3] Then for any other gas one needs to know *only* L_{lim}. This quantity, in turn, can be determined *only* by an *extrapolation* to $p = 0$ of values of L_p taken at various different pressures. The problem is then one of extrapolation, with all the uncertainty attending such a process. The uncertainty is of two kinds,—first that connected with the assumed functional form of L_p, and second that connected with the purely experimental errors.

Fortunately in the case of the permanent gases, the first cause of uncertainty is virtually absent. From work over a long range of pressures it is known that the curvature of L_p when plotted against p, between $p = 0$ and one atmosphere, is so small as to be quite negligible. We are therefore entirely justified in assuming a linear relation and this has been done in all recent work on these gases.

In the case of easily condensible gases, however, the plot of L_p against p sometimes shows a definite curvature. This makes the problem of extrapolation more difficult, and incidentally greatly increases the number of possible methods. Another addition to the list of possible methods comes from the fact that $1/L_p$ is often plotted against p. This custom arose in the following way. We see from Eq. (2) that pv is inversely proportional to L_p. Hence, except for the constant factor w, a plot of $1/L_p$ is also a plot of pv. If $w = 1$, the two become identical. We have

ourselves occasionally been confused by the use in the literature of pv for $1/L_p$, and in the present article we shall therefore write pv *only* when p and v are alone known (i.e., w unknown). We shall always write $1/L_p$ for the value of pv with $w = 1$.

Because of the reciprocal relation, Eq. (2), of pv and L_p, one gets the following useful relations, which also *define* λ and λ_0[4]

$$L/L_{\text{lim}} = (pv)_0/(pv)_1 = 1 + \lambda = 1/(1 - \lambda_0), \quad (4)$$

where L_{lim} and $(pv)_0$ refer to zero pressure, and L and $(pv)_1$ to unit pressure. In this article unit pressure is always one atmosphere. It would be much more logical to use L_1 and L_0 in place of L and L_{lim}, but the nomenclature of Eq. (4) is so well established (except that P and V, or p and V, are sometimes used for our p and v), that in order to facilitate comparison with the many articles quoted, we here adopt it. From Eq. (4) it follows immediately that

$$\lambda_0 = \lambda/(1 + \lambda) \quad \text{and} \quad \lambda = \lambda_0/(1 - \lambda_0). \quad (5)$$

Returning to the possible methods for determining L_{lim}, we then have two general classes. In the first class we plot L_p against p and extrapolate to get the intercept L_{lim} at $p = 0$. In the second class we plot $1/L_p$ against p and extrapolate to get the intercept $1/L_{\text{lim}}$ at $p = 0$. Each of these classes applies also to pv data. That is, if we have made observations only on p and v, then we may plot the products pv

---

[2] R. T. Birge, Rev. Mod. Phys. **1**, 1 (1929). See p. 12. (This article will be denoted *G.C.* 1929.)

[3] Thus $R = (M)_{O_2}/(L_{\text{lim}})_{O_2} = 32/(L_{\text{lim}})_{O_2}$.

---

[4] The quantity λ was introduced originally in connection with a linear function of p, and is particularly appropriate in this case. It is, however, now used for the more general case, and this has led to certain misconceptions that make it desirable to emphasize its correct definition, as given by Eq. (4).

The particular symbol λ is in general use, except among the Germans, who write $1 - \alpha$ in place of $1 + \lambda$ (see Geiger und Scheel, Handb. d. Physik **2**, 493 (1926)). In *G.C.* 1929 the German nomenclature was adopted and this choice has since been criticized by Batuecas, J. chim. phys. **29**, 26 (1932). The symbol λ_0 is being introduced for the first time, in place of the standard symbol A_0^1, for reasons that will appear in connection with Eq. (57) ahead. Meanwhile we wish to emphasize the fact that in this article λ and λ_0 are defined always and only by Eq. (4), although the actual numerical values obtained for λ and λ_0, from a given set of data, will naturally depend on how the data are treated.

against p, and extrapolate to get the intercept $(pv)_0$. This is, except for the constant factor w, precisely the same as plotting $1/L_p$ against p. Similarly we may plot $1/pv$ against p, which is equivalent to plotting L_p.

If values of L_p had been determined in all investigations, there would be no need to discuss pv values at all. Actually, however, much density work has been done at atmospheric pressure only. This yields the quantity L. Then in order to get $L_{\lim}$ it is necessary only to get values of pv at or near one atmosphere, and at various lower pressures. This is often done in a separate investigation. From these values of pv one obtains, by any one of a variety of methods, $(pv)_0$ and $(pv)_1$. Their ratio by Eq. (4) gives $1+\lambda$. Then, also by Eq. (4)

$$L_{\lim}=L/(1+\lambda) \qquad (4')$$

and by Eq. (3)

$$M=LR/(1+\lambda). \qquad (6)$$

This last equation is the standard form used in the literature, and often, as we have said, the value of L comes from *one* investigation, or set of investigations, and the value of $(1+\lambda)$ from *another* set. We do wish, however, to emphasize the fact that in Eq. (6) $L/(1+\lambda)$ is only a long way of writing the $L_{\lim}$ of Eq. (3) and that the problem of determining M, aside from the value of R, is *solely* that of determining $L_{\lim}$. If values of L_p have been obtained, then $L_{\lim}$ follows directly by extrapolation, and it only obscures the issue to introduce separate values of L and $1+\lambda$. However, in order to compare our results with those given in the literature, we shall indicate how L and $1+\lambda$, as well as $L_{\lim}$, may be obtained from the data.

The importance of the above facts is so great that we here summarize the four general ways in which the data may be handled. In all four cases what is desired is a $f(p)$ and this is obtained most reliably by the method of least squares. By $f(p)$ we mean, however, the adopted analytic representation of the data, regardless of how obtained.

(1) One derives $L_p=f(p)$ from observed values of L_p. The value of $f(p)$ at $p=0$ is $L_{\lim}$, and this completes the problem.

(2) One derives $1/pv=f(p)$ from observed values of pv. The value of $f(p)$ at $p=0$ is $1/(pv)_0$ and at $p=1$ is $1/(pv)_1$. The ratio gives, by Eq. (4), the value of $1+\lambda$. This is all that can be derived from pv data.

(3) One derives $1/L_p=f(p)$. The value of $f(p)$ at $p=0$ is $1/L_{\lim}$ and this completes the problem.

(4) One derives $pv=f(p)$, and as in case (2) the values of $f(p)$ at $p=1$ and $p=0$ give $1+\lambda$, by Eq. (4).

Since L_p and $1/pv$ differ by only a constant factor, whatever is the best *functional* form of $f(p)$ in case (1) is necessarily the best in case (2), for any given gas. The same thing is true of (3) and (4). On the other hand, the best function for (1) and (2) may be different from that for (3) and (4) and regardless of whether it is different or not, the *numerical* value of $L_{\lim}$ derived from (1) will never, except by accident, agree with that derived from (3). If, however, one has set up a really satisfactory $f(p)$ in both (1) and (3), the difference in the two resulting values of $L_{\lim}$ will be found entirely negligible. Exactly similar remarks apply to (2) and (4).

We now come to the much more involved and uncertain problem of the derivation of a suitable $f(p)$, in any one of our four cases. With the exception of an arbitrary graphical extrapolation of the data, by means of some curved line, all methods are equivalent to the assumption of some particular functional form. Of all such possible forms, only two seem to have been used, *viz.*, the first and second degree rational integral functions (r.i.f.) commonly called linear and parabolic solutions. We shall consider specific forms for each of these two functions, and noting again that cases (1) and (2) necessarily require the same function, and similarly (3) and (4), we have the following new four-fold division.

(a) L_p or $1/pv$ parabolic
(b) L_p or $1/pv$ linear
(c) $1/L_p$ or pv parabolic
(d) $1/L_p$ or pv linear

Case (a)

We begin with (a), the parabolic equation for L_p, *viz.*,

$$L_p=a+bp+cp^2 \qquad (7)$$

where, by definition $a=L_{\lim}$. If now a value of a is obtained by least squares, or by some other

reliable method, this completes the problem. The following discussion, which is unfortunately somewhat involved, is necessary because L_p data have quite regularly been used, in various different ways, to calculate also values of λ or λ_0, and it seems desirable to consider the relative reliability of the values so obtained.

To obtain λ or λ_0 it is necessary to evaluate a, b, and c, and as usual we recommend for this a least squares solution. Then λ and λ_0 are given by Eq. (4), in which *both* L and L_{lim} represent *calculated* values. Very often in the literature, however, the *observed* value of L has been used, and in order to distinguish the two resulting values of λ or λ_0 thus obtained, we shall for the remainder of this section write L_a and L_c for the *adopted* (observed) and *calculated* values of L, respectively.[5] Then, by definition

$$L_c = a + b + c \qquad (8)$$

and as noted $L_{\text{lim}} = a$. Hence by Eq. (4)

$$1 - \lambda_0 = 1/(1 + \lambda) = L_{\text{lim}}/L_c = a/(a + b + c) \qquad (9),$$

or

$$\lambda_0 = (b + c)/(a + b + c) \quad \text{and} \quad \lambda = (b + c)/a. \qquad (10)$$

Eq. (10) gives the only correct values of λ_0 and λ corresponding to Eq. (7). There is also a simple geometrical relation between λ_0 and Eq. (7) which can be shown as follows. Let us *define*

$$\lambda_p' = (1/L_c)(dL_p/dp) = (b + 2cp)/(a + b + c). \qquad (11)$$

λ_p' can be thought of as the variable slope of the curve obtained by plotting L_p/L_c against p. The limiting value of this slope, at $p = 0$, is

$$\lambda' = b/(a + b + c), \qquad (12)$$

and this is not simply related to either λ or λ_0, as is seen by comparing Eqs. (12) and (10). But the value of λ_p' at $p = 1/2$ is exactly λ_0, by Eqs. (11) and (10). This and other relations are

---

[5] In the general Eq. (4) any value of L can be used, but it is obvious that the best value of λ or λ_0 will be obtained only when one uses the best calculated value of $L(=L_c)$. When, as is often the case, the value of L is known more accurately than the values at lower pressures, a proper weighting of the various points will obviously lead to a calculated value of $L(=L_c)$ close to the observed value $(=L_a)$. The same remarks apply to $(pv)_1$.
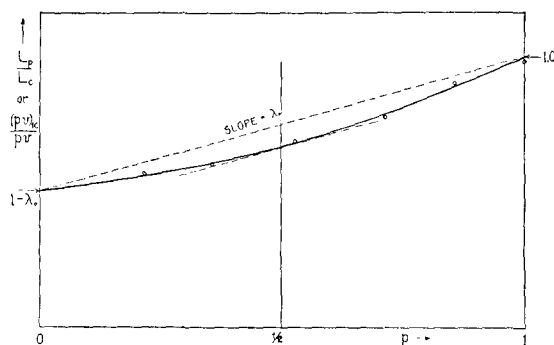

FIG. 1. Schematic diagram, applying to case (a), see Eqs. (11), (12), (24) and footnote 6.

shown in Fig. 1.[6] In place of the λ_p' of Eq. (11) we might have defined another variable slope, obtained by replacing L_c by L_{lim}. In that case the value of this new variable, at $p = 1/2$, is exactly λ. (See Fig. 2.) This interchange of λ_0 and λ, when L and L_{lim} are interchanged, is very general.
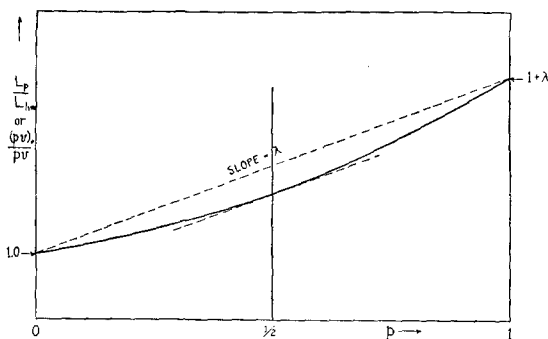
FIG. 2. Similar to Fig. 1, except that L_{lim} replaces L_c (or $(pv)_0$ replaces $(pv)_{1c}$), and λ replaces λ_0. In both figures the intercept values at $p = 0$ and at $p = 1$ atmosphere are indicated.

Our chief purpose in writing Eq. (11) is to make a comparison with the following Eq. (13), which has been used in the literature. In Eq. (13) the derivative of Eq. (11) is replaced by a special type of finite difference, and for convenience L_a is used in place of L_c. We then have, by definition,

$$\lambda_\Delta' = (1/L_a)(L_a - L_p)/(p_1 - p). \qquad (13)$$

---

[6] It is easy to show from Eq. (4) that λ_0 is the slope of the secant of the L_p/L_c : p curve, drawn between $p = 1$ and 0, and the slope of the secant equals that of a parabola at the middle point.

If now the values of λ_Δ' are considered as a $f(p)$ and are extrapolated by any method to $p=0$, one gets for this limiting value, remembering that $p_1 =$ one atmosphere.

$$\lambda_0 = (L_a - L_{1\text{im}})/L_a$$
or
$$1 - \lambda_0 = L_{1\text{im}}/L_a = 1/(1+\lambda). \tag{14}$$

This method then gives a value of λ_0 or of λ, but it is *not* the best value, even if the best possible extrapolation is made, since it uses an arbitrarily adopted value of L, ordinarily the observed value, in place of the best calculated value.

When a nonlinear equation is assumed in gas density work, it is quite customary to reduce it to a dimensionless form, by dividing the $f(p)$ by some common factor. Such a method, however, merely complicates the matter, so far as an evaluation of $L_{1\text{im}}$ is concerned, and leads to a value of λ_0 or λ suffering from the same defect as the value given by Eq. (14). This may be shown as follows. The customary procedure is to divide each observed value of L_p by the observed value of L. We then have, in place of Eq. (7)

$$L_p/L_a = a'' + b''p + c''p^2. \tag{15}$$

From Eqs. (7) and (15) we get

$$L_a = a/a'' = b/b'' = c/c''. \tag{16}$$

The limiting value $L_{1\text{im}}/L_a$ at $p=0$, or a'', is commonly used to get λ or λ_0 as follows:

$$L_{1\text{im}}/L_a = a'' = 1 - \lambda_0 = 1/(1+\lambda). \tag{17}$$

The desired value of $L_{1\text{im}}$ is evidently $L_a a''$, and it is quite immaterial what numerical value of L_a is used in Eq. (15), so long as the *same* value is used in getting $L_{1\text{im}} = L_a a''$. Moreover, the value of $L_{1\text{im}}$ so obtained is identical with the a of Eq. (7). This is shown directly by Eq. (16). The division and subsequent multiplication by L_a is thus entirely useless, so far as the calculation of $L_{1\text{im}}$ is concerned. On the other hand, the value of λ_0 and λ given by Eq. (17) suffers from the same defect as does that given by Eq. (14), i.e., the observed value of L has been assumed to be the best value. The correct expression for the best value of λ_0 or λ resulting from Eq. (15)

is obtained, as usual, by first putting $p=0$ and then $p=1$, and taking the ratio. This gives

$$L_{1\text{im}}/L_c = a''/(a'' + b'' + c'')$$
$$= 1 - \lambda_0 = 1/(1+\lambda). \tag{18}$$

This is identical with Eq. (9), because of the relations of Eq. (16). *If* now, by chance, $L_c = L_a$ we have, by Eq. (15), $a'' + b'' + c'' = 1$, and Eq. (18) becomes identical also with Eq. (17).[7]

Let us now consider the parabolic equation in $1/pv$, corresponding to Eq. (7), *viz.*

$$1/pv = a' + b'p + c'p^2. \tag{19}$$

To get the best value of λ or λ_0 we treat this equation just as we did Eq. (7), i.e., we evaluate

---

[7] There are a number of different possible values of λ and λ_0, derivable from the same set of data, that should be distinguished. Taking Eq. (4) as our basis, we note that the value of λ or λ_0 depends on, (1) the value adopted for $L_{1\text{im}}$, and (2) the value adopted for L. In Eq. (7) the value of $L_{1\text{im}}(=a)$, which as just noted is identical with that obtained from Eq. (17), corresponds to a parabola passed freely through the experimental values L_p, preferably by least squares. On the other hand, the value of $L_{1\text{im}}$ in Eq. (14), which comes in turn from Eq. (13), corresponds to some curve (it may be a parabola) that is *required* to pass through the adopted value L_a. Hence the $L_{1\text{im}}$ of Eq. (7) or (17) is preferable to that of Eq. (14). The λ or λ_0 derived from *either* Eq. (17) or (14) is, however, definitely inferior to that derived from Eq. (9) or (18), since in both Eqs. (17) and (14) we assume the adopted L_a, whereas in Eqs. (9) and (18) we assume the calculated L_c value given by Eqs. (7) and (15).

We note also that the non-dimensional L_p/L_a data may be treated by finite differences, just as in the case of Eq. (13). In fact, by merely putting L_a inside the parentheses, Eq. (13) becomes the appropriate equation. We see that the λ_0 of Eq. (14), expressed in this way, is merely the limiting value of the slope of a set of secants, drawn from the adopted $L_a/L_a = 1$ value at $p=1$ to the successive experimental L_p/L_a points. See Fig. 3.

Certain further remarks about Eq. (13) may appropriately be made here. It follows immediately from the form of Eq. (13) that if one assumes for λ_Δ' a *linear* equation in p, and gets a least squares solution, then the calculated values of L_p, or of L_p/L_a, that may be derived by Eq. (13) from the calculated values of λ_Δ', satisfy a *parabola*. This calculated parabola, as already noted, necessarily passes through the observed L_a point, but it is *not* the least squares parabolic solution of the L_p data under the condition that the parabola pass through the L_a point. In general if λ_Δ' is assumed as a r.i.f. of the rth degree, this is equivalent to assuming L_p to be a r.i.f. of the $(r+1)$th degree which is required to pass through a definite point.

a', b', and c' by least squares or otherwise. We again differentiate between an adopted and a calculated value of $(pv)_1$, by writing $(pv)_{1a}$ and $(pv)_{1c}$ respectively. Then, by Eq. (4),

$$1-\lambda_0 = 1/(1+\lambda) = (pv)_{1c}/(pv)_0$$
$$= a'/(a'+b'+c'). \quad (20)$$

This corresponds to Eq. (9), and is obviously in agreement with Eq. (9) since, from Eqs. (2), (7) and (19)

$$L_p(pv) = w = a/a' = b/b' = c/c'. \quad (21)$$

Corresponding to Eq. (10) we write, from Eq. (20)

$$\lambda_0 = (b'+c')/(a'+b'+c')$$
and $\hspace{7cm} (22)$
$$\lambda = (b'+c')/a'.$$

Corresponding to Eq. (11) we again get a simple geometrical relation by writing

$$\lambda_p' = (pv)_{1c}\frac{d(1/pv)}{dp} = \frac{b'+2c'p}{a'+b'+c'} = \frac{b+2cp}{a+b+c}. \quad (23)$$

The value of λ_p' at $p=1/2$ is, by Eq. (22), exactly λ_0. In this case λ_p' is the variable slope of the $(pv)_{1c}/pv$ curve, considered as a function of p, but this is identical with the L_p/L_c curve mentioned in connection with Eq. (11) and shown in Fig. 1, since, from the inverse proportionality of L_p and (pv), Eq. (2), we have

$$L_p/L_c = (pv)_{1c}/pv. \quad (24)$$

Accordingly both sides of Eq. (24) are given as the ordinate in Fig. 1. We may now define a new variable slope, by replacing $(pv)_{1c}$ by $(pv)_0$ in Eq. (23). This slope, at $p=1/2$, will give λ (see Fig. 2). We may also treat $1/pv$ data by finite differences, but as noted near the end of footnote 7, we are thus effectively treating the data in a non-dimensional form, and we accordingly take this up first.

Dividing each $(1/pv)$ by an adopted $(1/(pv)_{1a})$ value we get

$$(pv)_{1a}/(pv) = L_p/L_a = a''+b''p+c''p^2. \quad (25)$$

In other words, due to the inverse proportionality of L_p and pv, the non-dimensional Eqs. (15) and (25) are identical. In the case of pv data we cannot of course evaluate $L_{\lim}$, but the remarks

regarding the evaluation of λ or λ_0 from Eq. (15), by means of Eq. (17), apply also to the present case. To be specific,

$$(pv)_{1a}/(pv)_0 = L_{\lim}/L = a''$$
$$= 1-\lambda_0 = 1/(1+\lambda) \quad (17')$$

gives a poor value of λ_0 or λ. If Eq. (25) is properly handled, the value of $(pv)_0$ corresponds to a parabola passed freely through the pv data, but the value of λ_0 or λ in Eq. (17') suffers from the use of the arbitrarily adopted $(pv)_{1a}$ value. The correct equation is

$$(pv)_{1c}/(pv)_0 = L_{\lim}/L_c = a''/(a''+b''+c'')$$
$$= 1-\lambda_0 = 1/(1+\lambda). \quad (18')$$

It is quite common practice to handle the non-dimensional $(pv)_{1a}/pv$ data by means of finite differences.[8] In analogy to Eq. (13), remembering

---

[8] See A. Guye, J. chim. phys. **17**, 141 (1919), for an extended discussion of this and other special methods for handling the data. For a criticism of some of his conclusions see G. P. Baxter, J. Am. Chem. Soc. **44**, 595 (1922). It is well known that it is convenient to handle finite differences only when the abscissas are equally spaced, and the use of Eq. (26) in the literature is confined to this special case. We should like at this point to make some general remarks on the handling of finite differences. Let us suppose that a difference table shows that the successive second differences are constant, within the limits of experimental error. Then one is justified in adopting a parabolic formula. Common practice is now to take the *average* second difference and apply it to the final first difference and the final observation to get the first extrapolated value, and so on for further extrapolated values. In such a solution *only* the first two and last two observations are used. No other observations have any effect on the result. The parabola passing through each of the extrapolated points has the following properties. It passes exactly through the last two observations and lies on one side of the first two observations by some small but equal amount. This particular solution cannot be obtained by least squares with any set of weights but it does effectively give infinite weight to the last two observations, equal but finite weight to the first two, and zero weight to all the others. A special case is that of three equally spaced observations. In such a case the parabola necessarily passes through each of the three points and the question of weighting does not enter. Also the calculated and observed values of pv coincide and there is necessarily only one possible value of λ_0 or λ. Guye used the method only in this special case. The real question in such a situation is whether, with only three observations, there is any reason to assume a parabolic formula. The deviation from linearity may be due entirely

also the remark in footnote 7, we write

$$\lambda_\Delta' = [1 - (pv)_{1a}/\overline{pv}]/(p_1 - p). \qquad (26)$$

That this expression is identical with Eq. (13) is seen from the relation between $\overline{pv}$ and L_p given in Eq. (25). Hence the limiting value again gives λ_0. In this case λ_0 is the limiting value of a set of secants drawn on the $(pv)_{1a}/(\overline{pv})$ curve from unity at $p=1$ to the successive experimental points (compare final paragraph of footnote 7, and see Fig. 3).
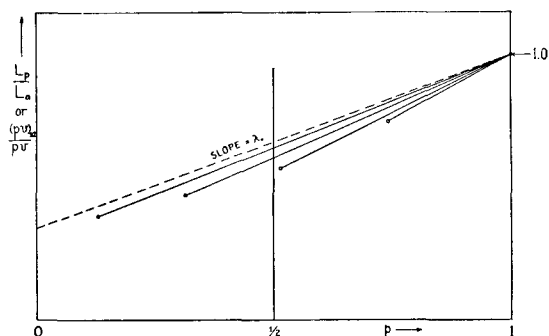

FIG. 3. When L_p/L_a is plotted against p, the limiting value of the slope of a set of secants, drawn from unity at $p=1$ to the experimental points is one (but not the best) estimate of λ_0 (see footnote 7). This is also true if $(pv)_{1a}/\overline{pv}$ is used as ordinate (see Eq. (26)).

This completes the discussion of case a, which may be summarized as follows. One fits data on L_p to a parabolic function, Eq. (7), and evaluates the coefficients a, b and c, preferably by least squares. Then the best value of the desired $L_{\lim}$ is given by the coefficient a, and the best value of λ or λ_0, if desired, by Eq. (10). λ_0 is geometrically the slope, at $p=1/2$, of the L_p/L_c plot, where L_c is the calculated value at $p=1$ given by Eq. (7). λ is similarly the slope, at $p=1/2$, of the $L_p/L_{\lim}$ plot. A non-dimensional form of the parabola, Eq. (15), leads by a longer process to the same value of $L_{\lim}$ as that given by Eq. (7), but as usually employed, to a less trust-

to experimental error. An example of this situation is given in the last part of the section on Illustrations.

In a similar way a set of first differences showing no trend justifies the use of a linear equation, but if one uses, for the extrapolation, the average first difference, this is equivalent to passing a straight line exactly through the first and last observations and ignoring all the others.

worthy value of λ or λ_0, Eq. (17), since an arbitrary value L_a is used in place of the calculated L_c. The best value is given by Eq. (18). A commonly employed finite difference method for handling the data, Eq. (13), not only uses L_a in calculating λ or λ_0, but forces the adopted curve, which may be a parabola, to pass through the adopted L_a point.

In the case of $1/\overline{pv}$ data one can calculate only λ or λ_0, the best values resulting from the adopted parabola, Eq. (19), being given by Eq. (20) or Eq. (22). The non-dimensional form, Eq. (25), is identical with Eq. (15), and leads to the correct value of λ or λ_0 given in Eq. (18') or to the poorer value of Eq. (17'). The finite difference method for handling such non-dimensional data, Eq. (26), although often employed, suffers from the same defects as Eq. (13). Some of the relations in the case of the non-dimensional forms are given in Figs. (1) to (3).

Case (b)

The linear case (b) is merely a special form of case (a), in which the coefficient of the second degree term in p has become negligibly small. We therefore *assume* this coefficient to be zero, and thus get a linear equation. Every equation in case (a) can be used for case (b) by setting $c = c' = c'' = 0$. Thus one takes Eq. (7) in the form

$$L_p = a + bp. \qquad (27)$$

The best value of $L_{\lim}$ is as before merely a, while the best value of λ_0 or λ is given by

$$1 - \lambda_0 = 1/(1+\lambda) = L_{\lim}/L_c = a/(a+b) \qquad (28)$$

or

$$\lambda_0 = b/(a+b) \quad \text{and} \quad \lambda = b/a. \qquad (29)$$

In place of a variable slope for the L_p/L_c plot, we now have a *constant* slope

$$\lambda_0 = (1/L_c)(dL_p/dp) = b/(a+b). \qquad (30)$$

We can use Fig. 1 also for this case, if we assume that the secant shown in that figure is now the best representation of the data.

In the case of the linear equation, the more general coefficients of Eq. (27) may conveniently be replaced by λ_0 and L_c. The result is

$$L_p = L_c[1 + \lambda_0(p-1)]. \qquad (31)$$

If in Eq. (30) we replace L_c by $L_{1\text{lim}}$ we get

$$\lambda = (1/L_{1\text{lim}})(dL_p/dp) = b/a, \qquad (32)$$

i.e., λ is the constant slope of the $L_p/L_{1\text{lim}}$ plot, and the secant of Fig. 2 can be taken to represent such a plot. Introducing λ and $L_{1\text{lim}}$ in Eq. (27) we get

$$L_p = L_{1\text{lim}}(1 + \lambda p). \qquad (33)$$

There is no particular advantage in using, for case (b), the finite difference Eq. (13), although one *might* adopt for λ_0 the *average* value of the various secant slopes (λ_Δ') thus obtained. Such a value of λ_0 has naturally the same defect discussed in connection with Eqs. (13) and (14).

We may also use the data in the non-dimensional form

$$L_p/L_a = a'' + b''p, \qquad (34)$$

but as in the case of Eqs. (15) and (17), the resulting value $L_{1\text{lim}} = a''L_a$ is identical with that given by Eq. (27), while the use of a'' as a value of $1 - \lambda_0$ is inadvisable since it assumes the correctness of the point L_a. The best value of λ_0 or λ is given by

$$L_{1\text{lim}}/L_c = a''/(a'' + b'') = 1 - \lambda_0 = 1/(1 + \lambda) \quad (35)$$

and Eqs. (35) and (28) are really identical, just as were Eqs. (18) and (9).

The linear equation in $1/pv$ may now be given a similar treatment. We begin with Eq. (19) in the form

$$1/pv = a' + b'p \qquad (36)$$

and note that a trustworthy value of λ_0 or λ can be obtained only by evaluating a' and b' by some good method. The result is

$$1 - \lambda_0 = 1/(1 + \lambda) = (pv)_{1c}/(pv)_0$$
$$= a'/(a' + b'), \quad (37)$$

which is really identical with Eq. (28), on account of the general relations given in Eq. (21). Hence

$$\lambda_0 = b'/(a' + b') \quad \text{and} \quad \lambda = b'/a'. \qquad (38)$$

By analogy with Eq. (30), we have in place of the variable slope of Eq. (23), the constant slope

$$\lambda_0 = (pv)_{1c}[d(1/pv)/dp] = b'/(a' + b') \qquad (39)$$

and the secant of Fig. 1 can be used to represent this constant slope of the $(pv)_{1c}/pv$ plot. Intro-

ducing λ_0 and $(pv)_{1c}$ into Eq. (36) we get

$$1/pv = (1/(pv)_{1c})[1 + \lambda_0(p - 1)]. \qquad (40)$$

If in Eq. (39) we replace $(pv)_{1c}$ by $(pv)_0$, the result is

$$\lambda = (pv)_0[d(1/pv)/dp] = b'/a' \qquad (41)$$

as shown by the secant of Fig. 2. Introducing λ and $(pv)_0$ into Eq. (36) one has

$$1/pv = [1/(pv)_0](1 + \lambda p). \qquad (42)$$

As already noted, it is rather common to use pv data in non-dimensional form, and in this linear case Eq. (25) becomes

$$(pv)_{1a}/pv = L_p/L_a = a'' + b''p, \qquad (43)$$

which is identical with Eq. (34). The remarks in connection with Eqs. (34) and (35) therefore apply here.

We note finally that Eqs. (30) and (39) are *special* definitions of λ_0, holding only for a linear equation, and similarly Eqs. (32) and (41) are special definitions of λ. These special definitions appear implicitly in the linear Eqs. (31), (33), (40) and (42).

Case (b) may be summarized as follows. One fits the data on L_p to a linear function, Eq. (27), and evaluates the coefficients a and b. The best value of the desired $L_{1\text{lim}}$ is given by a, and the best value of λ or λ_0, if desired, by Eq. (29). λ_0 is geometrically the constant slope of the L_p/L_c plot, and λ that of the $L_p/L_{1\text{lim}}$ plot, as shown explicitly by Eqs. (30) and (32) and implicitly by Eqs. (31) and (33). A non-dimensional form of the data, Eq. (34), is of no practical advantage, and is often used to give a poor value of $1 - \lambda_0$ ($= a''$) in place of the correct Eq. (35). In the case of $1/pv$ data, the linear Eq. (36) leads to values of λ_0 and λ given properly by Eq. (37) or Eq. (38). λ_0 is geometrically the constant slope of the $(pv)_{1c}/pv$ plot, and λ that of the $(pv)_0/(pv)$ plot, as shown explicitly by Eqs. (39) and (41), and implicitly by Eqs. (40) and (42). The secants of Figs. 1 and 2 also show these relations.

Case (c)

All of the necessary mathematical forms are included in cases (a) and (b). To get the analogous formulas for cases (c) and (d) one merely

substitutes $1/L_p$ for L_p and pv for $1/pv$. This change often results in the replacement of λ_0 by $-\lambda$ or of $1-\lambda_0$ by $1+\lambda$. Aside from changes of this type the equations of cases (a) and (c) are strictly analogous. The analogy holds also for the assumptions made and for the reliability of the results. Hence it is unnecessary to repeat the greater part of the discussion of case (a), and in general we merely list each equation of case (c), with the *analogous* equation of case (a) indicated in *square* brackets.

The basic parabolic equation is

$$1/L_p = \alpha + \beta p + \gamma p^2, \qquad (44) \ [7]$$

where, by definition, $\alpha = 1/L_{1\text{im}}$, and

$$1+\lambda = 1/(1-\lambda_0) = (1/L_{1\text{im}})/(1/L_c)$$
$$= \alpha/(\alpha+\beta+\gamma) \quad (45) \ [9]$$

or

$$\lambda = -(\beta+\gamma)/(\alpha+\beta+\gamma)$$

and
$$\qquad (46) \ [10]$$

$$\lambda_0 = -(\beta+\gamma)/\alpha.$$

For pv data the basic equation is

$$pv = \alpha' + \beta' p + \gamma' p^2. \qquad (47) \ [19]$$

We define

$$\lambda_p'' = L_c \frac{d(1/L_p)}{dp} = \frac{1}{(pv)_{1c}} \frac{d(pv)}{dp}$$

$$= \frac{\beta+2\gamma p}{\alpha+\beta+\gamma} = \frac{\beta'+2\gamma' p}{\alpha'+\beta'+\gamma'}. \quad (48) \ [11] \ [23]$$

Hence λ_p'' is the variable slope of the L_c/L_p or of the $pv/(pv)_{1c}$ plot, and by Eq. (46) equals $-\lambda$ at $p=1/2$. This and other relations are shown in Fig. 4, which is analogous to Fig. 1 of
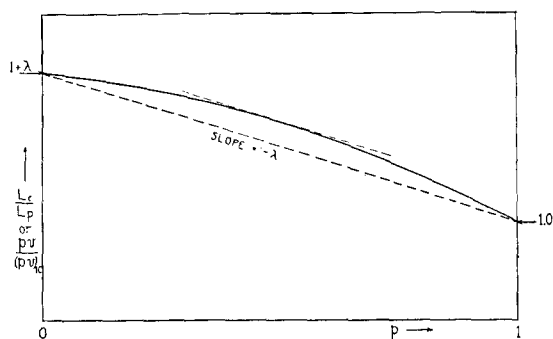

FIG. 4. Illustrative of case (c), Eq. (48), and strictly analogous to Fig. 1, of case (a).
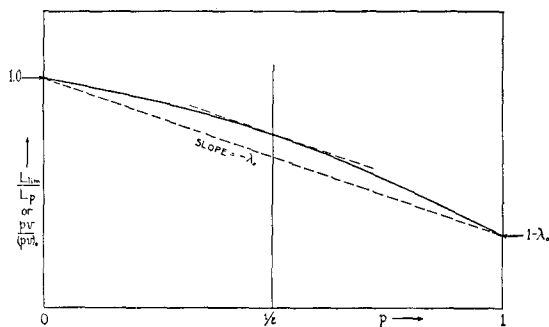

FIG. 5. Illustrative of case (c), and related to Fig. 4 in exactly the same way that Fig. 2 is related to Fig. 1, so far as the ordinate is concerned.

case (a). Replacing L_c by $L_{1\text{im}}$, or $(pv)_{1c}$ by $(pv)_0$, in Eq. (48) gives a variable slope with a value $-\lambda_0$ at $p=1/2$. This is shown in Fig. 5, which is analogous to Fig. 2.

In non-dimensional form we have

$$L_a/L_p = pv/(pv)_{1a}$$
$$= \alpha'' + \beta'' p + \gamma'' p^2 \quad (49) \ [15] \ [25]$$

and L_a/α'' gives a value of $L_{1\text{im}}$ identical with that obtained from Eq. (44). On the other hand, α'' is a poor value of $1+\lambda$ or of $1/(1-\lambda_0)$, (compare discussion of Eqs. (17) and (17')). The best value is given by

$$L_c/L_{1\text{im}} = (pv)_0/(pv)_{1c}$$
$$= \alpha''/(\alpha''+\beta''+\gamma'')$$
$$= 1+\lambda = 1/(1-\lambda_0). \quad (50) \ [18] \ [18']$$

A common use of the non-dimensional data is in the finite difference form

$$\lambda_\Delta'' = \frac{1-(L_a/L_p)}{p_1-p} = \frac{1-pv/(pv)_{1a}}{p_1-p}. \quad (51) \ [13] \ [26]$$

λ_Δ'' is the slope of the secant drawn from unity, at $p=1$, to any other experimental point, in terms of the L_a/L_p or $pv/(pv)_{1a}$ data, and its limiting value at $p=0$ gives a poor value of $-\lambda$, just as the limiting value of λ_Δ' of Eqs. (13) and (26) gives a poor value of λ_0. This is illustrated in Fig. 6.

Case (c) has been given in such condensed form that a summary seems unnecessary.

Case (d)

In a similar manner we write down, for case (d), equations analogous to those of case (b),
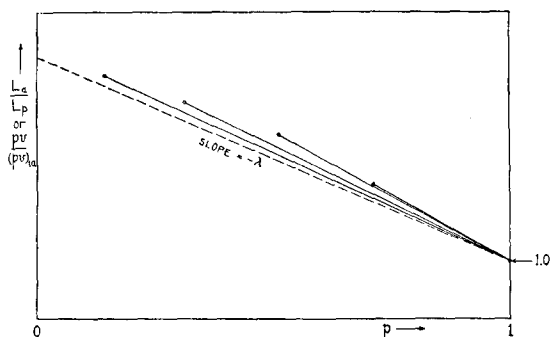
FIG. 6. Illustrative of case (c), Eq. (51), and strictly analogous to Fig. 3 of case (a).

starting with the assumed special form of Eq. (34), in which $\gamma = 0$.

$$1/L_p = \alpha + \beta p. \qquad (52)\ [27]$$

The best value of $1/L_{\mathrm{lim}}$ is α, and the best value of λ or λ_0 is given by

$$1 + \lambda = 1/(1 - \lambda_0) = \alpha/(\alpha + \beta) \qquad (53)\ [28]$$

or

$$\lambda = -\beta/(\alpha + \beta) \quad \text{and} \quad \lambda_0 = -\beta/\alpha. \qquad (54)\ [29]$$

For pv data we use the special form of Eq. (47),

$$pv = \alpha' + \beta' p, \qquad (55)\ [36]$$

giving

$$1 + \lambda = 1/(1 - \lambda_0) = \alpha'/(\alpha' + \beta'). \qquad (53')\ [37]$$

λ_p'' of Eq. (48) now equals $-\lambda$, the constant slope of the L_c/L_p or $pv/(pv)_{1c}$ plot, as given by the special definition

$$-\lambda = L_c \frac{d(1/L_p)}{dp} = \frac{1}{(pv)_{1c}} \frac{d(pv)}{dp}$$

$$= \beta/(\alpha + \beta) = \beta'/(\alpha' + \beta'). \qquad (56)\ [30]\ [39]$$

The secant of Fig. 4 can be taken to represent such a plot. If in Eq. (56) we replace L_c by L_{lim}, and $(pv)_{1c}$ by $(pv)_0$ we get

$$-\lambda_0 = L_{\mathrm{lim}} \frac{d(1/L_p)}{dp} = \frac{1}{(pv)_0} \frac{d(pv)}{dp}$$

$$= \beta/\alpha = \beta'/\alpha'. \qquad (57)\ [32]\ [41]$$

Hence $-\lambda_0$ is the constant slope of the L_{lim}/L_p or $pv/(pv)_0$ plot, which may be represented by the secant of Fig. 5. Eqs. (56) and (57) give

the best values of λ and λ_0, as obtained from the coefficients α' and β' of Eq. (55). We note also that Eq. (57) is the origin of our symbol λ_0, which bears the same relation to $(pv)_0$ in Eq. (57) that λ does to $(pv)_{1c}$ in Eq. (56).[9]

Putting L_c, L_{lim}, λ and λ_0 into Eq. (52) we get the special equations

$$1/L_p = (1/L_c)[1 - \lambda(p - 1)] \qquad (58)\ [31]$$

and

$$1/L_p = (1/L_{\mathrm{lim}})(1 - \lambda_0 p). \qquad (59)\ [33]$$

Putting $(pv)_{1c}$, $(pv)_0$, λ and λ_0 into Eq. (55) we get the special equations

$$pv = (pv)_{1c}[1 - \lambda(p - 1)] \qquad (60)\ [40]$$

and

$$pv = (pv)_0(1 - \lambda_0 p). \qquad (61)\ [42]$$

We may use the $1/L_p$ or pv data in the non-dimensional form

$$L_a/L_p = pv/(pv)_{1a} = \alpha'' + \beta'' p. \qquad (62)\ [34]\ [43]$$

As in Eq. (49), L_a/α'' gives a value of L_{lim} identical with that obtained from Eq. (52), and as in the case of Eq. (49), α'' is a poor value of $1 + \lambda$ or $1/1 - \lambda_0$. The best value is given, as in Eq. (50), by

$$L_c/L_{\mathrm{lim}} = (pv)_0/(pv)_{1c} = \alpha''/(\alpha'' + \beta'')$$

$$= 1 + \lambda = 1/(1 - \lambda_0). \qquad (63)\ [35]$$

There is no particular advantage in using, for case (d), the finite difference Eq. (51), although one might adopt for $-\lambda$ the average value of the various secant slopes (λ_Δ'') thus obtained. This value of $-\lambda$ corresponds to a linear graph forced to pass through an accepted point at $p = 1$.

This completes the discussion of the four cases, and we now turn to the literature for illustrations of the use of the more important of the preceding equations.

ILLUSTRATIONS

As our first illustration we consider the possible methods of handling a set of L_p data, and for this purpose we choose five values of L_p

---

[9] On this same ground λ_1 is a more logical symbol than λ, but we do not wish to depart from such a well established nomenclature.

for $(CH_3)Cl$, published by Batuecas.[10] These are listed in Table I.

TABLE I.

p (at.)	1	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{3}$	$\frac{1}{4}$
L_p (g. liter^{-1})	2.3075	2.2895	2.2802	2.2713	2.2664

Each of these L_p values is itself the average of many determinations, all of which appear in the original paper. Batuecas gives a plot of L_p as a $f(p)$. When replotted on a large scale, both L_p and $1/L_p$, as $f(p)$, show a small but perceptible deviation from linearity. This deviation may of course be due to some systematic experimental error, a question to which we revert later. For the moment we accept the data at their face value, and proceed to the analytic derivation of a value of $L_{1\text{lim}}$. In this work we use the method of least squares, giving equal weight to each point. In order to indicate certain relations, we have carried out the calculation for all four types of suggested solution, i.e., L_p and $1/L_p$ linear, and L_p and $1/L_p$ parabolic in p. In the following sections, unless otherwise stated, L always means the calculated value L_c.

(1) L_p linear. Eq. (27)

$$L_p = (2.25291 \pm 0.00012) + (0.05467 \pm 0.00019)p,$$

$s = 1.33 \times 10^{-4}$ or one part in 17,300 of L.[11]

Hence

$$L_{1\text{lim}} = a = 2.25291 \pm 0.00012, \quad L = a + b = 2.30758,$$

λ, from Eq. (4) or (29), $= 0.024266 \pm 0.000084$.

(2) $1/L_p$ linear. Eq. (52)

$$1/L_p = (0.443793 \pm 0.000031)$$
$$- (0.010454 \pm 0.000051)p,$$

$s = 0.353 \times 10^{-4}$ or one part in 12,300 of L.

---

[10] T. Batuecas, Anales soc. espān. fís. quím. **24**, 528 (1926).

[11] The standard deviation (S.D.) of a set of observations is $s = (\Sigma wv^2/\Sigma w)^{\frac{1}{2}}$ where $v =$ residual $= Obs.-Calc.$, and w is the weight of each observation. When all observations are given equal weight (unity) as here, s becomes $(\Sigma v^2/n)^{\frac{1}{2}}$, where n is the number of observations. The S.D. differs from the average deviation of the observations from their calculated values only in that the root mean square average is used. The probable error of each coefficient evaluated and of the function itself is proportional to s.

Hence

$$L_{1\text{lim}} = 1/\alpha = 2.25330 \pm 0.00016,$$

$$L = 1/(\alpha + \beta) = 2.30766,$$

λ, from Eq. (4) or (54), $= 0.02412 \pm 0.00012$.

(3) L_p parabolic. Eq. (7)[12]

$$L_p = (2.25249 \pm 0.00030) + 0.056316p$$
$$- 0.0013042p^2,$$

$s = 1.087 \times 10^{-4}$ or one part in 21,230 of L.

Hence

$$L_{1\text{lim}} = a = 2.25249 \pm 0.00030,$$

$$L = a + b + c = 2.30751,$$

λ, from Eq. (4) or (10), $= 0.024424$.

(4) $1/L_p$ parabolic. Eq. (44)

$$1/L_p = (0.443948 \pm 0.000056) - 0.011068p$$
$$+ 0.0004891p^2,$$

$s = 0.201 \times 10^{-4}$ or one part in 21,560 of L.

Hence

$$L_{1\text{lim}} = 1/\alpha = 2.25252 \pm 0.00028,$$

$$L = 1/(\alpha + \beta + \gamma) = 2.30750,$$

λ, from Eq. (4) or (46), $= 0.024408$.

The above data show the following significant relationships:

(a) The proportional value of s for the linear and parabolic solutions of L_p is not very different, and the two resulting values of $L_{1\text{lim}}$ agree as well as is to be expected.[13] On the other hand the proportional value of s for the linear solution of $1/L_p$ is distinctly larger, and the corresponding value of $L_{1\text{lim}}$ differs from the parabolic solution by more than is to be expected from the probable errors. Hence we conclude that the linear solution of the $1/L_p$ data is unsatisfactory.

(b) Both parabolic solutions lead to almost the same value of $L_{1\text{lim}}$ and to the same probable error. The latter results from the near equality of the two proportional values of s. We conclude

---

[12] In the case of parabolic solutions it has seemed unnecessary to calculate probable errors except for the important first coefficient giving $L_{1\text{lim}}$.

[13] Their difference just equals the sum of their probable errors.

therefore that either parabolic solution is satisfactory, and slightly better than the linear solution of L_p.

(c) In spite of the smaller proportional value of s for the parabolic solutions, as compared with the linear, the corresponding probable errors are much larger. This is an illustration of the fact that as the number of undetermined coefficients in the adopted function is increased, the reliability of any extrapolation very rapidly decreases.[14] We may elaborate this very important point as follows. In the region of *interpolation*, the probable error of any adopted function is a correct indication of the reliability. But in the region of *extrapolation* we are *assuming* the correctness of the adopted function, and any calculated probable error is based on such an assumption. Hence if the assumption is incorrect, the probable error will be quite misleading. So in the present case, the relatively small probable error of L_{lim} for the two linear solutions is a true criterion of the reliability of L_{lim} only if the function is known to be linear. In the case of $(CH_3)Cl$, however, it is possible that the function is not linear, and therefore the probable error of L_{lim}, as calculated by either parabolic solution, is a much more conservative indication of the reliability.

We accordingly adopt, as the best value of L_{lim} obtainable from the data of Table I, the average of the two parabolic solutions, i.e.,

$$L_{lim} = 2.2525 \pm 0.0003.$$

It will be shown in the second article of this series that the best value of R is 22.4146 ± 0.0003 liter-mole^{-1}. Hence by Eq. (3) $M = 50.4889 \pm 0.0067$ for $(CH_3)Cl$. The international atomic weights for H ($= 1.0078$) and Cl ($= 35.457$) seem well established and should have negligibly small errors compared to that in M itself. Hence $C = 12.0086 \pm 0.0067$ from the data of Batuecas.[10] In order to compare our results with the calculations of Batuecas, we note that the best value of λ, as obtained from the average of the two

parabolic solutions is 0.02442. All four types of solution give a value of L close to the observed 2.3075, and the parabolic solutions give this exactly, to five digits.

Batuecas makes the calculation in a quite different manner. He first plots L_p against p and makes a graphical linear extrapolation, obtaining $L_{lim} = 2.2527$, compared to our 2.2529 for the same function. Batuecas' article also includes a number of sets of measurements of pv at various pressures, the highest pressure being usually close to one atmosphere. These data have not been considered in our calculations. From such data he calculates λ_A'' of our Eq. (51).[15] The values so obtained are found to *decrease* rapidly with decreasing pressure. This corresponds to a pv curve with a negative (and large) coefficient of the second degree term, whereas his $1/L_p$ data lead, as we have just seen, to a *positive* (and small) second degree coefficient. The two sets of data are therefore definitely incompatible, and since the average deviation of the pv data is much larger than for the L_p data, we have discarded the former data entirely. Batuecas, on the other hand, considers that the decrease in the λ_A'' values is due to experimental difficulties (surface effects, etc.) and adopts for λ the *highest* λ_A'' value in each set. This happens to agree fairly well with his previous determination of λ. The average of all such highest values is 0.0245.

Batuecas now uses his observed value of L and his graphically extrapolated value of L_{lim} to get, by Eq. (4), $\lambda = 0.0243$. This he averages with his other value, 0.0245, and combines the new average λ with his observed L to get finally $L_{lim} = 2.25254$. By chance this is close to our adopted L_{lim} value, but we fail to see the justification for his handling of the pv data, or for his averaging of the results.[16]

---

[14] This fact has been well shown by H. Schultz, J. Am. Statistical Assoc. **25**, 139 (1930) (see plots on p. 159). This important article was unfortunately overlooked in a later discussion of the same subject by one of us (R. T. Birge, Phys. Rev. **40**, 207 (1932)).

[15] Batuecas calculates the ratio of the change of pv to that of p, using always the highest pressure (which is *not* one atmosphere) for one of the points. This ratio is then divided by $(pv)_1$, a value he evidently gets by some sort of extrapolation. His resulting λ_A'' values are therefore not exactly those defined by Eq. (51), and the limiting value at $p = 0$ is not exactly λ, but the theoretical discrepancy is negligible.

[16] In particular one is never justified in combining an observed value of L with a value of λ obtained from pv data in order to calculate L_{lim}, when L is one of a set of

In order to leave no false impression regarding the properties of $(CH_3)Cl$, we may note that later work on this gas has been performed by Crespi.[17] He obtains for L_p at $p = 1$, $1/2$ and $1/4$ atmospheres the values 2.30734, 2.27998, and 2.26617 respectively. These lie quite accurately on a linear graph, giving

$$L_p = (2.252489 \pm 0.000062)$$
$$+ (0.054868 \pm 0.000093)p,$$

$s = 0.43 \times 10^{-4}$ or one part in 53,600 of L.

Hence $L_{lim} = 2.252489 \pm 0.000062$,

$\lambda = 0.024359 \pm 0.000041$,

$L = 2.30736$,

$M = 50.4886 \pm 0.002$,

$C = 12.0082 \pm 0.002$.

Because of the small number of points, the stated probable errors are rather unreliable, but it would appear that Crespi's work is somewhat more accurate than that of Batuecas, and hence that the curvature in the L_p plot of Batuecas' data is due to some systematic experimental error. Fortunately both agree very well in the resulting values of L_{lim} and the atomic weight of carbon.

Moles,[18] in summarizing the work on $(CH_3)Cl$, handles the data of Batuecas[10] and Crespi[17] in an entirely different way. Although each of these investigations yields directly a value of L_{lim}, and as just noted, practically the same value, Moles does not attempt to get a value of L_{lim} from either. Instead he adopts Batuecas' own average value of λ (0.0245), and uses our finite difference Eq. (51) to get from Crespi's data a

value $\lambda = 0.0244$.[19] The average of these two ($= 0.02445$) is then combined with *Crespi's observed* value of L to obtain L_{lim}. The resulting value (2.25227) differs appreciably from the value 2.2525 given directly by each investigation. This well illustrates the danger of working with an observed value of L and a calculated value of λ, rather than with a directly calculated value of L_{lim}. Moles' resulting value $C = 12.003$ differs from ours solely because of the difference in the adopted values of L_{lim}.

It seems desirable to interpolate, at this point, a brief description of the method for calculating λ from (pv) data, ordinarily used by Batuecas and others who have been associated with Professor A. Guye. The method is described in detail in a paper by Batuecas, Maverick and Schlatter[20] on the value of $1 + \lambda$ for oxygen.

One obtains just three values of pv at different pressures, the first pressure being close to one atmosphere. These are fitted to a linear equation for pv—our Eq. (55)—using a and b for our α' and β'. The value of $1 + \lambda$ then follows from our Eq. (53'). The method used is really a variation of the finite difference method, Eq. (51), and is as follows. From points 1 and 2, using Eq. (55), β' is evaluated, i.e., $\beta' = \Delta(pv)/\Delta p$. From points 1 and 3 a second value of β' is obtained. The resulting average value of β' is substituted in Eq. (55) to give, for each experimental point, a value of α'. The three resulting values are averaged. Then our Eq. (53') is used, with $(\alpha'_{aver} + \beta'_{aver})$ in the denominator, but with *each* of the three values of α' in the numerator, to get three values of $1 + \lambda$, which are finally averaged. The last two steps are obviously unnecessary. The average $1 + \lambda$ is given directly by $\alpha'_{aver}/(\alpha'_{aver} + \beta'_{aver})$.

By making a graphical representation of this method one can see easily that it is equivalent to a least squares solution in which a high weight has been given to point 1 (at the highest pressure), a much smaller weight to point 2 and a still smaller weight to point 3. This is probably a more reasonable method of weighting than

---

L_p values. Such a procedure forces a slope on the L_p plot that may be quite incompatible with that obtained directly from the L_p data. In such a situation the only reliable L_{lim} value is that obtained directly from the L_p data. In the language of statistics, the value of λ obtained from pv data may be considered as *a priori* information regarding the value of λ to be obtained from the L_p data. A method for incorporating such information into the solution of the L_p data for L_{lim} could doubtless be devised, but the method would probably not be simple and so far as we know has never been investigated.

[17] Crespi, Anales soc. españ. fís. quím. **25**, 25 (1927).

[18] E. Moles, Rec. trav. chim. **48**, 864 (1929). The L_p data of Crespi[17] are listed here.

[19] The values of λ_Δ'' are assumed by Moles to vary linearly, and the correct extrapolated value at $p = 0$ is then actually 0.02444.

[20] T. Batuecas, G. Maverick and C. Schlatter, J. chim. phys. **22**, 131 (1925). See p. 137.

that of equal weights. Hence we have no real objection to the method, although we feel that it is preferable to make a deliberate weighting of each point rather than to use this concealed and wholly arbitrary system of weighting.[21]

We conclude this section with a discussion of a paper by Dietrichson, Orleman and Rubin,[22] entitled *The Density of Ammonia at Reduced Pressures and its Relation to the Atomic Weight of Nitrogen, the Gas Constant R, and the Limiting Molal Volume, v_0,* which illustrates the confusion on this subject that still remains in the literature and the consequent need of an extended treatment of the theory such as we have just given. These investigators determine the density, ρ_p, of ammonia at $p=1$, $2/3$ and $1/3$ atmospheres.[23] They list and recalculate also previous data[24] on ρ_p at $p=1$, $2/3$, $1/2$, and $1/3$ atmospheres. D.O.R. give values of pv for $w=1$. Hence by our Eq. (2) this is merely $1/L_p$ and in general we shall write $1/L_p$ where they write pv.[25] They also use d_n for our L, v_0 for our R, and R for the gas constant which we shall denote R_0. ($R=R_0T_0$, or $v_0=RT_0$ in their notation.)

D.O.R. first write our Eq. (6) in the form

$$L = M(1+\lambda)/R_0T_0$$

and remark that the accuracy with which M can be determined depends on the accuracy with which λ, R_0 and T_0 are known. We have, however, shown in detail that the value of λ is of only secondary importance. It is $L_{\lim}$ that is needed. But what is more important is that T_0 has nothing to do with the problem. It is only the product $R_0T_0=R$ that is required. This product

R is known far more accurately than either factor,[26] and as is shown in the next article of this series, has a probable error of only about one part in 64,000 ($R=22.4146\pm0.0003$). The final error in M is therefore due almost solely to the purely experimental error in $L_{\lim}$.

D.O.R. next describe three different ways in which λ may be calculated, from the same set of data, and label the three different numerical results thus obtained λ', λ'' and λ''' respectively. Their λ' is gotten from our Eq. (62), in which as we have remarked, α'' is a poor value of $1+\lambda$. To get their α'' they make a linear extrapolation of the L_a/L_p data to $p=0$. Their λ'' is gotten from our Eq. (34), in which again α'' is a poor value of $1/(1+\lambda)$. Here also they make a linear extrapolation of the L_p/L_a data.[27] Their first two methods then correspond to our cases (*d*) and (*b*) respectively, i.e., an assumption of $1/L_p$ and L_p linear in p, respectively.

In the third method, D.O.R. say "the authors feel that the coefficient of deviation from Boyle's law should be defined as $(1/pv)_{p\to0}[\partial(pv)/\partial(p)]_T = \lambda'''$." Since they use our Eq. (6) to get M, there is only one possible definition of $1+\lambda$, namely the one given in our Eq. (4). Their definition of λ''', as written, is ambiguous, but we find, by actual trial, that their published result ($\lambda'''=0.015109$) can be obtained *only* if one writes $\lambda''' = -[1/(pv)_0][d(pv)/dp]_{p=1}$. The true definition of λ, by the discussion following our Eq. (48), is however

$$-[1/(pv)_1][d(pv)/dp]_{p=1/2} = 0.014516.$$

These results are based on their published equation

$$pv = 1/L_p = 1.315403 - 0.017767p - 0.001054p^2.$$

---

[21] The question of weighting, and the actual experimental results of reference 20 will be discussed more fully in the second article of this series.

[22] G. Dietrichson, C. W. Orleman and C. Rubin, J. Am. Chem. Soc. **55**, 14 (1933). To be denoted D.O.R.

[23] Their actual results, which will often be used in this discussion, are $\rho_p=0.77126$, 0.51161, and 0.25458 respectively. These lead by our Eq. (2) to $L_p=0.77126$, 0.767415 and 0.76374 respectively, or $1/L_p=1.296580$, 1.303075 and 1.309346 respectively.

[24] E. Moles and T. Batuecas, Anales soc. espān. fís. quím **28**, 871 (1930). To be denoted M.B.

[25] They also use 1 cc as the unit of volume and therefore all their values of pv and coefficients in equations for pv as a $f(p)$ must be divided by 10^3 to obtain our $1/L_p$ in liter·atmos. g.$^{-1}$ units. We make such a division in quoting their results.

[26] Actually R_0 is determined from R/T_0 and has therefore practically the same probable error as has T_0, i.e., one part in 9100. ($T_0=273.18\pm0.03$ by *G.C.* 1929.)

[27] We have remarked in footnote 8 on the inadvisability of using any average first difference for such an extrapolation. Quite aside from this question, we are unable to check the results given by D.O.R. In the case of $1+\lambda'$ the two available first differences are 0.00501 and 0.004836, but to get their published 1.014901 one must assume 0.005055 in the extrapolation. Similarly for $1/(1+\lambda'')$ the available first differences are 0.004985 and 0.004765, but 0.005031 is needed to get their published 0.985219 ($=1/1.015003$).

In getting this equation the authors state that each point was weighted in proportion to the pressure and a least squares solution obtained. Since, however, there are three undetermined coefficients and only three points, the correct equation necessarily satisfies each point, regardless of the weighting, and could hardly be called a least squares solution. Their published equation gives a small positive residual for every point. The correct solution, which can be obtained very quickly from a difference table, is

$$1/L_p = 1.315390 - 0.017793p - 0.001017p^2.$$

Having obtained their three different values of λ, which lead in turn to three different values of M, the molecular weight of NH_3, the authors remark "It will be noted that the at.wt. of N . . . was found to show considerable variation. It therefore appears that the gas density method for the determination of atomic weights cannot at present be considered as reliable as that based on gravimetric analyses, not because of insufficient accuracy as regards the gas densities, but rather because of uncertainties in connection with λ, R_0 and T_0 (also $R = R_0 T_0$)." But the *variation* in the value of λ, as calculated in different ways, has nothing whatever to do with the value of R, and again we note that it is only the values of R and L_{lim} that concern us in this problem.

Having reached the above conclusion, D.O.R. now proceed to use our Eq. (3) to *calculate* R, from the accepted best value of M for NH_3 (17.0312). In this calculation they use, for the first time, the value of L_{lim} directly calculated from the data. This value they calculate in two ways, (1) $1/L_p$ parabolic, giving their $1/L_{lim} = 1.315403$ in the equation already quoted, and (2) $1/L_p$ linear, for which they get $1/L_{lim} = 1.316484$. From these two results they get $R = 22.4032$ and 22.4213, respectively. They make similar calculations for the M.B. data, getting $R = 22.4136$ and 22.4120. As we shall show immediately, the most probable value of L_{lim} obtainable from this work has a probable error of about one part in 12,000 and obviously cannot be used to improve the evaluation of R, which is already known to about one part in 64,000, from the work on oxygen to be discussed in the second paper of this series.

Let us now turn to a consideration of the most probable value of L_{lim}, and of M, resulting from the work of D.O.R. and of M.B. As just noted, D.O.R. have given solutions based on our Eqs. (44) and (52), i.e., $1/L_p$ parabolic and $1/L_p$ linear, respectively. We have made similar calculations, with results somewhat different from those that they publish. We have already given our parabolic solution for the D.O.R. data. The corresponding linear solution with the weighting (proportional to pressure) used by D.O.R., is

$$1/L_p = (1.31582 \pm 0.00011)$$
$$- (0.01922 \pm 0.00011)p,$$

$s = 0.50 \times 10^{-4}$ or one part in 26,000 of L.

$L_{lim} = 0.759983 \pm 0.000064$ or one part in 12,000.

λ (from Eq. (53)) $= 0.01482 \pm 0.00010$,

$M = 17.0347 \pm 0.0014,$ $N = 14.0113 \pm 0.0014.$

For the reasons to be stated presently, we believe this to be the best result obtainable from the D.O.R. data. Their own equation is

$$1/L_p = 1.316484 - 0.020066p$$

and gives $s = 2.17 \times 10^{-4}$, or 4.3 times that of the necessarily minimum value of the least squares solution.

For the M.B. data we obtain as the linear solution

$$1/L_p = (1.315776 \pm 0.000088)$$
$$- (0.01991 \pm 0.00012)p.$$

$s = 0.60 \times 10^{-4}$ or one part in 21,600 of L.

$L_{lim} = 0.760007 \pm 0.000051,$

$\lambda = 0.015371 \pm 0.000089,$

$M = 17.0352 \pm 0.0012,$ $N = 14.0118 \pm 0.0012.$

D.O.R. give for the same treatment of the M.B. data

$$1/L_p = 1.315940 - 0.020177p.$$

We find from this equation, $s = 0.96 \times 10^{-4}$, or 1.6 times the true least squares result.

Now a plot of the M.B. data shows no trend whatsoever from linearity. In fact an unweighted

parabolic solution of their data[28] gives $L_{1\text{im}}$ = 1.315846, which differs from the linear result by less than the probable error of the latter. On the other hand, the value of s for the D.O.R. linear solution is even smaller than that for the M.B. data, so that unless the former data are by far the more accurate, the deviation from linearity should be ascribed to experimental error. Moreover, the value $1/L_{1\text{im}} = 1.315390$, as already given for the parabolic solution of the D.O.R. data, leads to $M = 17.0403$, a result greater than the accepted value, 17.0312, by one part in 1870, whereas the linear solution gives a result greater by only one part in 4860. Finally, the linear solutions of the D.O.R. and the M.B. data give almost the same value of $L_{1\text{im}}$. We conclude, therefore, that the mean deviation of the D.O.R. data from a linear curve, i.e., one part in 26,000, is to be ascribed to experimental error. We are not in a position to discuss critically the cause of the high resulting value of M. We may note, however, that if the D.O.R. point at $p = 1/3$ is discarded, and a straight line passed through the other two $1/L_p$ values (at $p = 1$ and 2/3), the result is $L_{1\text{im}} = 0.759841$, or $M = 17.0315$ and $N = 14.0081$. It is therefore possible that systematic errors enter into the determination at $p = 1/3$.[29]

The value of $L_{1\text{im}}$, as published by M.B.[24] is 0.75990, as contrasted with our 0.760007 obtained from a linear $1/L_p$ solution. We have made solutions of their data corresponding to all four of our cases (a) to (d), but none of the results is as low as theirs.[30] From it, with their usual $R = 22.414$, M.B. obtain $N = 14.009$. It is possible that M.B. obtained their result in a manner similar to that just mentioned in con-

nection with the D.O.R. work, for if one uses only the two $1/L_p$ points at $p = 1$ and 2/3, the resulting value of $L_{1\text{im}}$ is 0.759886, giving, with our value of R, $N = 14.0091$. So again there may be systematic errors at their two lower pressures (1/2 and 1/3 atmospheres). In any case it would appear that ammonia is not a suitable material to use in precision measurements of the at.wt. of nitrogen, by the limiting gas density method.

Conclusions

From the preceding discussion we may conclude that, starting with the fundamental Eq. (3), the best value of $L_{1\text{im}}$ is obtained from the intercept of the L_p curve, at $p = 0$, after fitting the L_p data to an analytic function by means of least squares. In the case of the permanent gases, a linear function in p is known to be satisfactory, in the region $p = 1$ to 0 atmospheres, but in the case of the easily condensible gases, it may be necessary to use a parabolic function. Further, if L_p is given by a linear function in p, $1/L_p$ is likely to be an equally satisfactory linear function in p, and this form has often been used in the literature. A similar remark applies to L_p and $1/L_p$ as parabolic functions in p. These four possible functions have been treated in detail in cases (a) to (d). In each of these cases there are several ways in which the function may be written and handled, thus accounting for the wide variety of methods used in the literature. In each of the cases (a) to (d) however, we have shown that there is only one method that leads directly to the most reliable value of $L_{1\text{im}}$. Unfortunately some of the more commonly used methods of calculating $L_{1\text{im}}$ are by no means the most reliable, and numerical examples of this have been given in the concluding section on Illustrations.

If one has only values of pv, at various pressures, it is possible to calculate only λ or λ_0, as defined by Eq. (4). To get $L_{1\text{im}}$ we then need a value for L, the density at one atmosphere, as shown by Eq. (4'). It is, however, questionable procedure to use a value of L obtained from one investigation in combination with a value of λ, or λ_0, obtained in some other investigation. Furthermore, if experimental values of L_p are available, these lead directly to a value of $L_{1\text{im}}$, and the calculation of L and $1+\lambda$ from such

---

[28] We find that the weighted and unweighted linear solutions give values of $L_{1\text{im}}$ differing by only one part in 200,000, so that the weighting is here immaterial. All our parabolic solutions have accordingly, to save time, been made with unweighted data.

[29] The observed $\rho_n = 1/3$ is given by D.O.R. as 0.25458 ±0.00001 g. liter^{-1} and the calculated value, on this new assumption, is 0.254536. The discrepancy is 4.4 times the stated probable error, but the data quoted on p. 16 of D.O.R. indicate the possibility of systematic errors of this magnitude.

[30] This value, 0.75990, is quoted by the Inter. At. Wt. Committee, J. Am. Chem. Soc. **53**, 1627 (1931). We do not have available the original M.B. paper.

data, as is often done, only complicates the matter, and has sometimes resulted finally in an untrustworthy value of $L_{\lim}$. These points have also been discussed in connection with cases (a) to (d), and have been illustrated by examples from the literature. It is, in fact, easily possible to obtain a value of $L_{\lim}$, and hence a molecular weight, that differs by several times its probable error from the most trustworthy value given by the data. Hence for data of the accuracy now obtainable in precision work on gas densities, the best analytic methods are not only suitable, but demanded. The actual application of such methods to the data on oxygen and on carbon will be given in detail in the next two papers of this series.