

RESEARCH PAPER RP934

Part of Journal of Research of the National Bureau of Standards, Volume 17,
November 1936

ACCURATE REPRESENTATION OF REFRACTIVE INDEX OF DISTILLED WATER AS A FUNCTION OF WAVE LENGTH

By Leroy W. Tilton

ABSTRACT

Technique in using the Ketteler-Helmholtz dispersion equation is discussed and validity in representing the dispersion of water over the visible range of wave lengths is tested by least-squares adjustments of sixth-decimal-place refractive indices. It is concluded that four parameters are sufficient for accuracy within the spectral range investigated.

CONTENTS

| | Page |
|------------------------------------------------------------------------|------|
| I. Introductory discussion of dispersion equations..... | 639 |
| II. Refractive-index data..... | 642 |
| III. Computational procedures..... | 643 |
| 1. Equation 1, two Sellmeier terms, four (independent) parameters..... | 644 |
| 2. Equation 4, infrared term expanded, five parameters..... | 646 |
| 3. Equation 11, both Sellmeier terms expanded, six parameters..... | 646 |
| 4. Equation 4, term in λ^4 omitted, four parameters..... | 646 |
| IV. Comparative merits of four, five, and six parameters..... | 647 |
| 1. Relative goodness of fit and Chi-test..... | 647 |
| 2. Concluding discussion..... | 649 |

I. INTRODUCTORY DISCUSSION OF DISPERSION EQUATIONS

Many dispersion formulas in use for representing the refractive indices of transparent media are approximations to the form

$$n^2 = a^2_{\infty} - \sum \frac{M}{L^2 - \lambda^2} + \sum \frac{m}{\lambda^2 - l^2}, \quad (1)$$

where

$$a^2_{\infty} = 1 + \sum \frac{M}{L^2} + \sum \frac{m}{l^2}$$

is the dielectric constant, L^2 and l^2 with appropriate subscripts are used to designate, respectively, the effective wave lengths of absorption bands in the infrared and the ultraviolet regions; and similarly M and m are corresponding parameters. This is the equation to which the complete Ketteler-Helmholtz equations reduce for regions exclusive of the absorption bands. Strictly speaking, vacuum values should be used for n , the index of refraction, and for λ , the wave

length, but the index of refraction of air is so near unity and its dispersion so small that dispersion equations are frequently used for the combined system air plus the medium in question. This practice has been followed in this paper and in evaluating parameters all wave lengths are expressed in microns.

In some cases it has been found that only one infrared term and one ultraviolet term are necessary for satisfactory approximations. With only these two Sellmeier terms there are in reality but four independent parameters to be determined and for a simultaneous solution the equation may be written as

$$n^2 = -\frac{n^2-1}{\lambda^4}(l^2L^2) + \frac{n^2-1}{\lambda^2}(l^2+L^2) - \frac{1}{\lambda^2}\left(\frac{Ml^2}{L^2} + \frac{mL^2}{l^2}\right) + \left(1 + \frac{M}{L^2} + \frac{m}{l^2}\right), \quad (2)$$

where, inclosed in parentheses, there is a linearly independent set of four groups of unknowns. If, then, equation 1 is limited to three terms, its five parameters may be computed from the four groups of parameters as formed in equation 2. Equation 2, however, is not necessarily suitable for direct adjustments by least squares because it is not solved¹ for n^2 . Thus errors in the observed values of n^2 are involved in the coefficients to be used in the least-squares process and this may or may not be of importance in the adjustment that is obtained.

A thoroughly reliable adjustment of the parameters of equation 1 can be made after an approximate solution is obtained and the process of differentiation applied for obtaining the equation

$$\Delta n^2 = \frac{\lambda^2}{\lambda^2-l^2}\left(\Delta \frac{m}{l^2}\right) - \frac{\lambda^2}{L^2-\lambda^2}\left(\Delta \frac{M}{L^2}\right) + \frac{m\lambda^2}{l^2[\lambda^2-l^2]^2}(\Delta l^2) + \frac{M\lambda^2}{L^2[L^2-\lambda^2]^2}(\Delta L^2), \quad (3)$$

which can then be used in adjusting by least squares the system of betterments inclosed in parentheses. These betterments in turn permit revision of the parameters which were used in the approximate solution. In applying a procedure of this kind it should be remembered that second and higher order derivatives are ignored. Consequently, it is important, perhaps especially so with this particular equation (see section III and fig. 1), that the approximate solution be so good that equation 3 can be used with validity.

It is evident that equation 1 is not desirably convenient for curve fitting and usually when two Sellmeier terms are to be used, at least one of them is expanded and a more convenient form such as

$$n^2 = a^2 - k\lambda^2 - p\lambda^4 - \dots + \frac{m}{\lambda^2-l^2} \quad (4)$$

is obtained. The remarkable success that has, in general, been attained in using this equation (even without the term in λ^4) is probably ascribable to the fact that not only one Sellmeier term but many are simultaneously approximated. That is,

¹ The fact that an equation is given for n^2 rather than for n is a relatively unimportant matter in its use for curve fitting. A least-squares adjustment of n^2 rather than of n is in effect an implicit weighting of the observations but the percentage variation in these weights over the total range of n is usually small and negligible. Throughout this paper it is assumed that errors in λ are small as compared with those otherwise involved in index measurement.

$$\left. \begin{aligned} a^2 &= a^2_{\infty} - \sum \frac{M}{L^2}, \\ k &= \sum \frac{M}{L^4}, \\ p &= \sum \frac{M}{L^6}, \text{ etc.}, \\ m, \text{ and } l^2 \end{aligned} \right\} (5)$$

are some of the potential values of the parameters in the expanded form. Incidentally this view shows, also, why it is unsafe to assume ² that the ratio k/p equals L^2 and gives the location of an effective absorption band.

Equation 4 may, for convenience in simultaneous solutions, be written

$$n^2 = \frac{n^2}{\lambda^2}(l^2) + \frac{1}{\lambda^2}(m - a^2l^2) - \lambda^2(k - pl^2) - \lambda^4(p) + (a^2 + kl^2), \quad (6)$$

but again, as for equation 2, a direct adjustment by least squares is, theoretically at least, inadvisable because the equation is not solved for n^2 . One may, of course, use the differential process and the equation

$$\Delta n^2 = \Delta a^2 - \lambda^2 \Delta k - \lambda^4 \Delta p + \frac{1}{\lambda^2 - l^2} \Delta m + \frac{m}{(\lambda^2 - l^2)^2} \Delta l^2, \quad (7)$$

but there is a more direct procedure ³ which the writer has tried and found satisfactory. If for l^2 an approximate value l_0^2 can be found from previous experience, or in any way whatsoever, then after setting

$$\left. \begin{aligned} \mu^2 &= \lambda^2 - l_0^2, \\ \epsilon^2 &= l^2 - l_0^2, \\ b^2 &= (a^2 - l_0^2 k - l_0^4 p), \\ h &= (k + 2l_0^2 p), \end{aligned} \right\} (8)$$

and

one may write

$$n^2 = b^2 - h\mu^2 - p\mu^4 + \frac{m}{\mu^2 - \epsilon^2} \quad (9)$$

to correspond to equation 4. Then the equation

$$n^2 = \frac{n^2}{\mu^2} \epsilon^2 + \frac{1}{\mu^2} (m - b^2 \epsilon^2) - \mu^2 (h - p \epsilon^2) - \mu^4 p + (b^2 + h \epsilon^2) \quad (10)$$

corresponds to equation 6 and can be used for a direct least-squares adjustment of the linearly independent set of unknowns ϵ^2 , $(m - b^2 \epsilon^2)$, $(h - p \epsilon^2)$, p , and $(b^2 + h \epsilon^2)$. If l_0^2 of equations 8 has been suitably chosen, then there will be no appreciable error introduced by neglecting in equation 10 the slight error in the coefficient, n^2/μ^2 , of the first right-hand term.

² See, for example, p. 390 of Wood's Physical Optics, 1911 edition.

³ Suggested to the writer by Dr. L. B. Tuckerman of this Bureau.

Referring again to equation 1, if both types of Sellmeier terms are expanded, there results

$$n^2 = a^2 - k\lambda^2 - p\lambda^4 - q\lambda^6 - \dots + \frac{v}{\lambda^2} + \frac{w}{\lambda^4} + \dots, \quad (11)$$

where, in addition to the system expressed by equation 5, $v = \Sigma m$, $w = \Sigma ml^2$, etc. This equation 11 is linear in the parameters and comparatively convenient for a direct least-squares solution. If data with several significant figures must be adjusted and if, as is often the case with these dispersion equations, there is serious loss of significance in the process of solving the normal equations, then it may be necessary or advisable to adjust betterments by use of the equation

$$\Delta n^2 = \Delta a^2 - \lambda^2 \Delta k - \lambda^4 \Delta p - \lambda^6 \Delta q - \dots + \frac{1}{\lambda^2} \Delta v + \frac{1}{\lambda^4} \Delta w + \dots, \quad (12)$$

which is obtained after differentiating equation 11.

II. REFRACTIVE-INDEX DATA

During a period of preparation for making definitive measurements on the index of refraction of distilled water, preliminary data were obtained by measurements on 24 samples, of approximately 9 ml each, taken successively from storage in a glass bottle of 5 gallons capacity on 24 working days between March 23 and May 4, 1931. The refractive indices of these samples were measured at 20.00° C for 25 different wave lengths from 3889 to 7679 angstroms. When measuring a single sample, only four or five spectral lines were used, but these were always so selected that each day's measurements were well distributed over the visible spectrum.

All index measurements were made by the method of minimum deviation using a water-jacketed hollow prism mounted in a stirred air bath on the table of a spectrometer. Temperatures were controlled within approximately $\pm 0.01^\circ$ C and were measured with a platinum-resistance thermometer. All index results were corrected to refer to dry air at 20.0° C and a pressure of 760 mm of mercury. The temperature-controlled prism housing and other auxiliary apparatus have been described in a previous paper,⁴ and a description of the hollow prism will be given in a forth coming paper with Mr. John K. Taylor, who cooperated in taking the data that are used in these tests of dispersion equations.

Previously published tests of this kind have usually been made with four- and five-decimal-place data. In the present case it was estimated that errors in the observed indices were confined to the sixth decimal place and perhaps were less than ± 5 parts per million.⁵ Moreover, these indices were preliminary to experiments which were to be more carefully conducted. The latter, however, were to be based on fewer spectral lines and to be confined within a somewhat more restricted total range in the wave-length interval. Consequently, for subsequent use in adjusting the definitive data, it seemed desirable to use first the preliminary results in testing and selecting

⁴ J. Research NBS 17, 389 (1936) RP919.

⁵ Many precautionary details necessary for refractive-index determinations of this order of accuracy have been discussed by the writer in former papers. For a brief summary and references see J. Research NBS 14, 417 (1935) RP776.

a dispersion formula for which the added systematic errors of a computational nature should, if possible, be confined approximately to the seventh decimal place. Schönrock's⁶ experience in accurately representing the refractive indices of fluorite with a five-parameter equation (4) indicated that success could probably be attained. Nevertheless, comparative tests were desired in order to select a convenient equation having a minimum number of parameters; also to permit an objective quantitative estimate of the systematic errors involved in its use for this particular medium, water.

For these reasons equations 1 (with three terms only), 4 (both with and without the λ^4 term), and 11 have been considered and tested in choosing a dispersion equation for use in adjusting, representing, and interpolating precise data on the refractive index of distilled water.

III. COMPUTATIONAL PROCEDURES

The numerical difficulties in making least-squares solutions for data having seven significant figures are formidable. Mere arithmetical accuracy in such computations on many data is difficult. In addition, during the process of solving normal equations,⁷ there sometimes occurs a loss of significant figures that may be troublesome even when only two or three digits are required for expressing the parameters that are to be adjusted. Theoretically, of course, this can be avoided by some suitable change in the initial equation or in the manner in which the solution is attempted. Practically, however, it is usually quicker and easier to resolve the normal equations,⁸ considering their coefficients as exact numbers and retaining throughout a second solution as many extra digits as were lost during the first attempt. When six or seven significant figures are required in some of the parameters it is evident, therefore, that the situation may become serious because the theoretically possible alternate procedure may be hopelessly involved and obscure, especially if one is concerned with as many as five or six unknowns.

It is, perhaps, partly for the reasons outlined above that dispersion parameters are seldom thoroughly and completely adjusted when four or more parameters are used. Usually a value based on entirely extraneous evidence is arbitrarily preassigned to one or more of the parameters. Empirically this course is usually justifiable and it is often advisable when the sole purpose is adjustment of the observations. It should be remembered, however, that this procedure may decidedly influence the values which will be obtained for some of

⁶ Z. Instrf. 40, 93-96 (1920); 41, 103-105 (1921).

⁷ Some of the numerical difficulties here discussed are, of course, not peculiar to normal equations but may occur in other computations. Dr. Tuckerman called the writer's attention to this simple illustrative example. Assume a current I flowing through a galvanometer of resistance R_g with a shunt R_s , where R_s is small in comparison with R_g , say approximately $1/10000$ of R_g . Let each of these three quantities be observed to three or four significant figures. Then it appears at first impossible to get three or four significant figures when computing $I_g = I - I_s$ where $I_s = \frac{IR_g}{R_g + R_s}$. Nevertheless, by considering the observed values as exact and retaining throughout the computation a sufficient number of significant figures, perhaps seven or eight, one obtains the same result (to three or four figures) that is found by using the more appropriate formula $I_g = \frac{IR_s}{R_g + R_s}$ and retaining only three or four significant figures in the computation.

⁸ There seems to be no necessity of revising the "observational equations" and the squares and cross products after they have once been correctly computed and written with a sufficient number of significant figures to yield, by the process of summation, normal-equation coefficients having as many significant figures as are found in the initial data, if the solution is direct, or in the residuals if the process be one of differential betterment.

the other parameters. Such arbitrary selection of one parameter may, therefore, vitiate or call in question the inferences which in some instances are drawn from adjusted values of the others.⁹ This point is particularly pertinent for dispersion equations of the type here considered. Some of these parameters are especially sensitive to arbitrary changes in others and yet their "adjusted" values are often regarded as having informative value regarding the location of resonance frequencies, or absorption bands.

1. EQUATION 1, TWO SELLMIEIER TERMS, FOUR (INDEPENDENT) PARAMETERS

In using equations 2 and 3 for adjusting four parameters in (three terms of) equation 1, the computational difficulties were somewhat aggravated by the unsuitable nature or form of this two-Sellmeier-term equation as a means of accurately representing the dispersion of water. That this equation, as thus limited, might be found relatively unsuitable could have been predicted from dispersion theory and the known large dielectric constant of water. However, it does not follow that this equation 1 must necessarily be inferior to equation 4 when the latter is limited to four terms.

A simultaneous solution with equation 2 gave obviously large residuals. Differential betterment by least squares with equation 3 effected such large changes in the parameters that the results were even worse than the initial computations. A second simultaneous solution with equation 2 (using a second set of observed indices) likewise resulted in computed indices that agreed poorly with the observations, so poorly that a continuation with betterments by equation 3 was not advisable. Thus it seemed imperative to obtain better approximate values before second derivatives could safely be ignored. For this purpose a direct least-squares solution with equation 2 was attempted, but a solution of the normal equations, using nine significant figures, failed to yield even one significant figure for the group $\left(\frac{M^2}{L^2} + \frac{mL^2}{l^2}\right)$ of unknowns.

In most instances of difficult solutions at least some significant figures are obtained in solving for one of the unknowns and one knows at once how many digits must be retained in a successful re-solution of the normal equations. In this case, however, there was no indication as to what computational precision would be sufficient. Consequently, using the normal equations, the three other groups of unknowns were then expressed in terms of $\left(\frac{M^2}{L^2} + \frac{mL^2}{l^2}\right)$ and this system was repeatedly solved for several even (exact) values of L^2 such as 1, 2, 4, 9, and 36, all of which correspond roughly with known absorbing frequencies in the infrared spectrum of water.

In every case within this wide range of L^2 these sets of parameters were computed with internal consistency to about nine significant figures and they satisfied the normal equations to about eight or nine significant figures. Of these various sets of parameters none seemed really satisfactory but the one with $L^2=2$ gave the lowest sum of the squared residuals, and $L^2=4$ was next best. Then after trying 2.1,

⁹ See, for example, E. Flatow, *Ann. Physik* [4] **12**, 93 (1903).

2.25, 2.4, and 3.0 and plotting (see fig. 1) the corresponding values of Σr^2 , a value of $L^2=2.25$ was selected as approximately an optimum. The corresponding set of parameters was then bettered by a least-squares solution with the aid of equation 3. Some significant figures were lost in this last process, but the betterments themselves were comparatively small. In the resulting dispersion formula,

$$n^2 = 1.7726479 - \frac{0.0315734}{2.2535795 - \lambda^2} + \frac{0.00701841}{\lambda^2 - 0.0092513}, \quad (13)$$

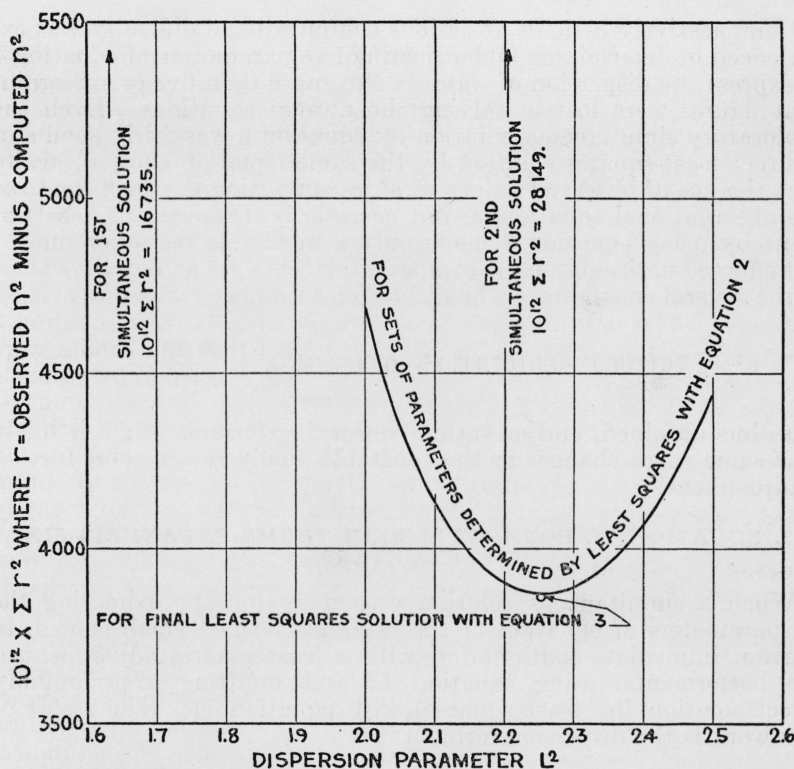


FIGURE 1.—Determining by trial the first of four dispersion parameters of equation 1.

In this difficult case of adjustment by least squares it was necessary to express three parameters in terms of a fourth and then plot Σr^2 for various arbitrarily selected values of the fourth. All sets of parameters having L^2 in the range 1 to 36 satisfied the normal equations to 8 or 9 significant figures.

the consistency of the full set of constants is adjusted with a precision fully as great as is warranted by the observations, but it is recognized that further least-squares adjustment by renewed application of equation 3 would introduce compensatory changes among these constants that in some instances might materially change the penultimate digits.

Incidentally, the final Σr^2 , although slightly lower than the similar summation for $L^2=2.25$ is not noticeably off the curve which has been plotted in figure 1 for those various computations that are all in accord (to eight significant figures) with normal equations obtained by the aid of equation 2. This means that errors in the observed values of

n^2 , which were allowed to enter in the right-hand of equation 2, were not in this instance of any importance. It also means that at least one of the constants of formula 13 can, with approximately the same validity in representing the observations, be rounded to three significant figures, provided it be considered as exact whenever used (including most particularly its use in the computation of the other constants in such manner that all of them form a consistent set that satisfies the normal equations).

2. EQUATION 4, INFRARED TERM EXPANDED, FIVE PARAMETERS

Comparatively little or no serious computational difficulty was experienced in determining and adjusting five parameters of equation 4 to express the dispersion of water. Not more than five or six significant figures were lost in solving the normal equations. Even the exploratory simultaneous solution by equation 6 was fairly good and a direct least-squares solution by the same equation showed, again, that the use of observed values of n^2 in computing the coefficients of the observational equations is not necessarily dangerous. A betterment by using equation 7 made only a negligible reduction in Σr^2 but effected noticeably large compensatory changes among the values of the several constants. The dispersion formula

$$n^2 = 1.7611069 - 0.0107941\lambda^2 - 0.0009048\lambda^4 + \frac{0.00653385}{\lambda^2 - 0.0140700} \quad (14)$$

was thus obtained, and as stated concerning formula 13, it is likely that some small changes in the constants would result upon further readjustment.

3. EQUATION 11, BOTH SELLMIEIER TERMS EXPANDED, SIX PARAMETERS

When a simultaneous solution was made for approximating the six parameters of equation 11 the residuals seemed small enough to warrant immediate continuation with a least-squares adjustment of the betterments, using equation 12 and omitting a preliminary direct solution (by least squares) with equation 11. The result of this work is the dispersion formula

$$n^2 = 1.7606279 - 0.0079050\lambda^2 - 0.0066381\lambda^4 + 0.0038681\lambda^6 + \frac{0.00649123}{\lambda^2} + \frac{0.000111406}{\lambda^4}, \quad (15)$$

for which Σr^2 is appreciably lower than for the initial (simultaneous) determination. Again, however, five or six significant figures were lost in solving the normal equations and some of the betterments were so large that the numerical coefficients in formula 15 are, perhaps, not in strictly optimum adjustment.

4. EQUATION 4, TERM IN λ^4 OMITTED, FOUR PARAMETERS

Equation 4, without the term in λ^4 , is one of the most commonly used equations for expressing dispersion. For simultaneous solutions four parameters may be easily evaluated from the system of equations:

$$\left. \begin{aligned}
 l^2 &= \frac{A\lambda_4^2 - \lambda_1^2}{A-1}, \\
 m &= \delta_2\delta_3\delta_4 \left(\frac{\Delta_{23} - \Delta_{34}}{\lambda_4^2 - \lambda_2^2} \right), \\
 k &= \Delta_{12} - \frac{m}{\delta_1\delta_2}, \\
 \text{and } a^2 &= n_1^2 + k\lambda_1^2 - \frac{m}{\delta_1},
 \end{aligned} \right\} (16)$$

where $A = \frac{\lambda_3^2 - \lambda_1^2}{(\Delta_{12} - \Delta_{23})} \left(\frac{\Delta_{23} - \Delta_{34}}{\lambda_4^2 - \lambda_2^2} \right)$; $\Delta_{12} = \frac{n_1^2 - n_2^2}{\lambda_2^2 - \lambda_1^2}$, $\Delta_{23} = \frac{n_2^2 - n_3^2}{\lambda_3^2 - \lambda_2^2}$, etc.;

$\delta_1 = (\lambda_1^2 - l^2)$, $\delta_2 = (\lambda_2^2 - l^2)$, etc., and the subscripts refer successively to values of n^2 and λ^2 from the four sets of observations that are to be used.

In using equations 16 it is obviously important to have the sets of observed data separated by wide intervals in wave length. For water, with observations extending over the visible spectrum, no serious loss of significance need occur, and the constants as thus obtained are often sufficiently reliable to serve as a useful basis for differential betterment by four terms of equation 7.

Direct least-squares solutions for these constants by using equation 6 also have proved fairly satisfactory for adjustment of the observations, but for a single least-squares process, where the initial results are not to be bettered, it is preferable to use equation 10. In general, however, whether using equations 7, 6, or 10, about four significant figures are lost in solving the normal equations. Therefore, if one desires to adjust the constants themselves it is unsafe to neglect a betterment based on the residuals that are given after any initial least-squares adjustment. From the particular data considered in this paper the dispersion formula

$$n^2 = 1.7616034 - 0.0119397\lambda^2 + \frac{0.00644940}{\lambda^2 - 0.0148669} \quad (17)$$

was finally obtained for comparison with formula 13 and the more elaborate formulas 14 and 15.

IV. COMPARATIVE MERITS OF FOUR, FIVE, AND SIX PARAMETERS

1. RELATIVE GOODNESS OF FIT AND CHI-TEST

When only one formula is used for representing a given set of data, some idea of its suitability for the purpose can be obtained by an examination of the residuals. Also, one may use the Chi-test¹⁰ to measure quantitatively in terms of probability the goodness of fit which is obtained, provided a numerical estimate of the precision of the data is known a priori. When, however, for an identical set of data, and identical assumptions concerning the weighting, there are several different formulas fitted by least squares to the observations, then these formulas may be directly compared as to relative goodness of fit by forming in each case an estimate of some characteristic error for an observation of unit weight. Estimates of this sort may be

¹⁰ For a recent discussion of this test, see W. Edwards Deming, *J. Am. Statistical Assn.* **29**, 372-382 (1934).

based on the assumption that the chances are about even for obtaining, purely by chance and exclusive of peculiarities or defects in the formula itself, a worse fit than has been found. Then, if probable error be the characteristic measure of precision selected for the basis of comparison it is estimated as $P. E. = 0.6745\sqrt{\Sigma r^2/C}$, where C is the number of observations decreased by the number of adjustable parameters in the formula.

The results of tests and intercomparisons of formulas 17, 13, 14, and 15 are given in table 1 where many of the entries are self-explanatory. Excepting for the unsuitable formula 13, none of the estimated probable errors seems unreasonably large, especially when it is considered that these data were preliminary observations and not made with the approved technique which was later adopted for the definitive measurements. Moreover, it should be noticed that these estimated probable errors do not differ for the equations 17, 14, and 15 (with four, five, and six parameters, respectively) by as much as the probable error in these estimates, namely, $0.477 P. E./\sqrt{C}$. Consequently, there is in this particular little or no definite indication that these three formulas would have the same relative standing if again compared by using them to represent another set of similar index data.

There should, of course, for this single set of data, be only one final estimate of the probable error of a single observation of weight unity. From the estimates in table 1 (and imperfection in a formula tends to make these values high) it appears that the probable error is not greater than $\pm 1.4 \times 10^{-6}$ in index and from the details of the observations the probable error is directly computed as $\pm 1.26 \times 10^{-6}$. Assuming, then, a probable error of $\pm 1.3 \times 10^{-6}$ as an a priori estimate, there have been computed from the Chi-tables, and listed in table 1 the probabilities, P , that the actually observed degrees of failure to fit might be worse purely by chance even if the corresponding formulas were perfectly suitable for the purpose of expressing the dispersion of water. For formula 17, then, the tabulated probability 0.23 is not seriously below the optimum value which is 0.50 approximately. In other words, even a perfect formula would, in about 23 of 100 such tests, fail to appear as suitable as does formula 17 when judged by this one test.

TABLE 1.—*Statistical comparison of dispersion formulas*

| | | | | |
|---------------------------------------------------------------------------------------------------------------------------------------|------------|------------|------------|------------|
| Designation of formula | 17 | 13 | 14 | 15 |
| Number of independent parameters | 4 | 4 | 5 | 6 |
| Number of + residuals (observed minus computed index) | 10 | 11 | 12 | 12 |
| Number of - residuals | 15 | 14 | 13 | 13 |
| Number of changes in sign of adjacent residuals | 12 | 6 | 11 | 14 |
| Number of non-changes in sign of adjacent residuals | 12 | 18 | 13 | 10 |
| $10^6 \Sigma r$ = algebraic sum of residuals | -0.03 | +0.04 | +0.18 | 0.00 |
| $10^6 \bar{x}$ = algebraic average residual | -0.001 | +0.002 | +0.007 | 0.000 |
| $10^6 \Sigma r $ = arithmetic sum of residuals | 36.2 | 94.1 | 35.4 | 34.3 |
| $10^6 \bar{x}$ = arithmetic average residual | 1.45 | 3.76 | 1.42 | 1.37 |
| $10^6 \bar{m}$ = median residual | 1.24 | 3.9 | 1.12 | 1.19 |
| $10^{12} \times \Sigma r^2$ | 94.6 | 541.4 | 84.0 | 72.9 |
| $10^6 \times$ estimated $P. E.$ (assuming that existing degree of fit can be worse by chance alone in 50 percent of such tests) | ± 1.43 | ± 3.42 | ± 1.38 | ± 1.32 |
| $10^6 \times$ estimated $P. E.$ of estimated $P. E.$ | ± 0.15 | ± 0.36 | ± 0.15 | ± 0.14 |
| Number of observations minus number of parameters | 21 | 21 | 20 | 19 |
| $\chi^2 = \Sigma r^2 / (1.483 P. E.)^2$, where $P. E.$ is estimated a priori as $\pm 1.3 \times 10^{-6}$ | 25.5 | 146 | 22.6 | 19.6 |
| Probability of worse fit by chance alone if $P. E.$ is estimated a priori as $\pm 1.3 \times 10^{-6}$ in index | 0.23 | Very small | 0.31 | 0.42 |
| Odds that formula is imperfect in form | 3.4 to 1 | Very great | 2.2 to 1 | 1.4 to 1 |

2. CONCLUDING DISCUSSION

The manner in which the four dispersion formulas fit the refractive-index data is illustrated in figure 2, where everything is referred to a datum or reference line, $\Delta n=0$, that corresponds to values of the refractive index of distilled water as computed by the general interpolation formula which has been adjusted to fit all approved data¹¹ on water taken in the refractometric laboratory at this Bureau. The four-constant formula 13 is obviously unsuitable and necessarily inaccurate for representing the data, and it is thus indicated that there is no single effective location of an absorption band in the infrared region that will serve to replace the bands which actually exist for water.

Referring to the average location and trend of the other curves in figure 2, the general lack of parallelism to the line $\Delta n=0$ is probably attributable to the accidental distribution of errors in this sample of

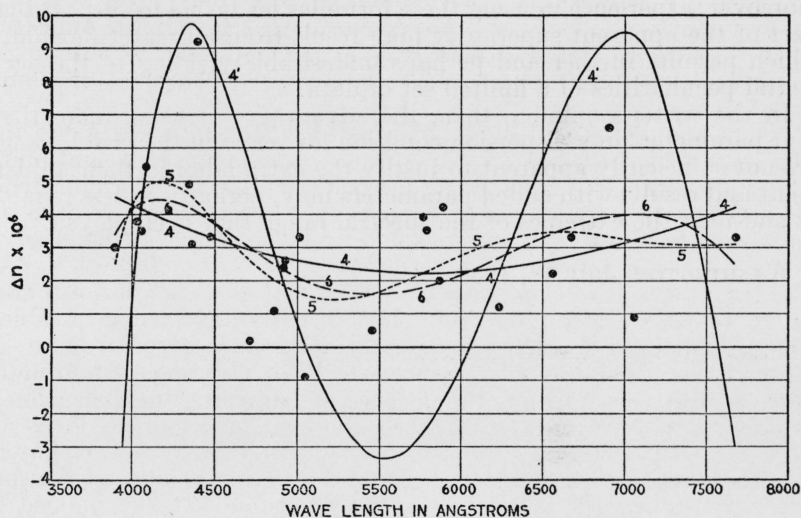


FIGURE 2.—Comparison of dispersion formulas.

Formulas 17, 13, 14, and 15, with 4, 4', 5, and 6 parameters, respectively, were adjusted by least squares to fit preliminary refractive-index data taken at 20.0° C on distilled water. Circular dots represent observations and lines represent computed indices. The reference line $\Delta n=0$ corresponds to indices computed by a general interpolation formula (based on formula 17) to fit more numerous data on distilled water which were subsequently taken with improved technique and procedures.

only 25 observations. One notices, particularly, an especially high index for the faint helium line 4388 Å; also the low value for the faint helium line 5048 Å, and the high value for the faint mercury line 6907 Å. The average increase of 3×10^{-6} in index which these samples show is attributed to slight contamination of the water which was not freshly distilled but had been stored in a dark-room where ammonia and other chemicals were kept. Moreover, during some of these initial experiments the hollow prism was almost entirely filled and a paper collar on the glass stem of the platinum-resistance thermometer was inadvertently allowed to touch the water. For these reasons,

¹¹ A total of 133 observations were used in adjusting by least squares a formula with 13 constants which represents the refractive index of distilled water in the visible region of the spectrum for the range 0 to 60° C. An account of this work by the writer, in collaboration with Mr. John K. Taylor, is being prepared for publication in this journal.

and also because complete tables of the refractive index and dispersion of water are to be included in a later account of work with more careful and approved procedures, no explicit values for these data are included in this paper.

It is evident from figure 2 that each of formulas 17, 14, and 15 agrees well with the average for all three of them. Aside from extremes of wave length, there is only one instance for which an individually computed index differs from this average by as much as 6×10^{-7} . Thus it seems that the Ketteler-Helmholtz dispersion formula 17 with four constants is probably adequate in representing the refractivity of water in the visible region to approximately $\pm 5 \times 10^{-7}$ or better.

As shown by certain data in table 1, there is, from a purely statistical viewpoint, some indication of slight added superiority of the five- and six-constant formulas for fitting this particular sample of 25 indices. On the other hand, the slight advantage is not reliably indicated and, moreover, experience in using these formulas leads one to suspect that part of the apparent superiority may result from a greater flexibility which permits further and perhaps undesirable yielding to the accidental peculiarities of a limited set of data.

In the writer's opinion, then, the advantages of using more than four parameters in a dispersion equation for water in the visible range are not sufficiently apparent to justify the extra labor in their adjustment and results with added parameters may, perhaps, be less reliable at and near the extremes of the spectral range that is employed.

WASHINGTON, July 18, 1936.