Pilgrims' progress in search of the fundamental constants

Jesse W. M. DuMond

Citation: Physics Today **18**(10), 26 (1965); doi: 10.1063/1.3046944
View online: http://dx.doi.org/10.1063/1.3046944
View Table of Contents:
http://scitation.aip.org/content/aip/magazine/physicstoday/18/10?ver=pdfcov
Published by the AIP Publishing

PILGRIMS' PROGRESS IN SEARCH OF THE FUNDAMENTAL CONSTANTS

The fundamental constants of nature are so interrelated that a measurement affecting one affects them all. The author became interested when Millikan's oil-drop value of the electron charge was different from the value given by x-ray determination of crystal spacings. To assist in finding the true values, he invented a method for plotting various functions of the constants in a space of as many coordinates as there are constants. If all measurements are consistent, the plotted functions intersect in a point. When they do not intersect, one examines standard deviations, which correspond to thicknesses of surfaces, in an effort to find out what is wrong. In three decades, searches of this kind have reduced uncertainties in the constants from a fraction of a percent to, at most, tens of parts per million.

By Jesse W. M. DuMond

The practice of making broadly inclusive surveys, from time to time, of the status of our knowledge of the fundamental constants of physics and chemistry may be said to have started with a famous paper by Raymond T. Birge, of Berkeley, published in *Reviews of Modern Physics* in 1929. To Professor Birge, also, is due the credit for being the first, as far as I know, to apply the method of least squares in order to determine most probable values of three of the constants; e, the electronic charge m, the electron rest mass; and h, Planck's constant, using a highly overdetermined set of experimental data on functions of these three quantities.

Since the fundamental constants are intricately interconnected, with many accepted theoretical interrelationships tying them together, Birge emphasized that the only appropriate way of achieving a consistent set of values of the constants requires application of the method of least squares. By "consistent" is meant a set of values satisfying all the theoretical relationships accepted as correct, at the epoch of a given adjustment, and simultaneously doing as little violence as possible (in the sense of least squares) to a carefully selected list of all the most precise measurements available at that time. At any given time, however, experience shows that some data are notably inconsistent with the consensus. This idea of consistency is one of enormous importance. It is, in fact, the only test the physical sciences have that they are on the right track and this, to my mind, is the main reason for studying the constants. To know a consistent set of approximate numerical values is useful, but to study the data critically and test whether the entire picture puzzle really fits together and makes sense, and if not, to point out where troubles lie, this, it seems to me, is both fascinating and crucially important and should be the true objective of the constants analyst.

I became interested in the constants only a few years later than Professor Birge when discussion was raging over the apparent discrepancy between R. A. Millikan's oil-drop value of the electron charge, e, and the value implied by x-ray determinations of the interplanar atomic lattice spacings of crystals. Compton and Doan in this country, and Thibaud in France, had just succeeded in determining, by diffraction from artificial ruled gratings, the wavelengths of certain x-ray emission lines in absolute cgs units. Using the same lines diffracted from crystal lattices, the absolute dimensions of the unit cell of the crystal structure, and hence its volume in cubic centimeters, could be calculated and, from the macroscopically measured density of the crystal, one could then calculate the mass of the unit cell. Comparison of this with the gram-molecular weight of the unit cell gave the Avogadro number, N, and the quotient of the electrochemically determined faraday by N gave

the value of e. The value so obtained turned out to be almost 0.6 percent higher than the oil-drop value. The work of Joyce A. Bearden,[1] in the United States, and of Bäcklin,[2] Söderman,[3] Tyrén,[4] and others, in Sweden, later pushed this method of normalization of the scale of x-ray emission-line wavelengths by means of ruled-grating diffraction to a precision of a part in ten thousand, or better.

I was a graduate student at Caltech when this controversy started. Our chief, Dr. Millikan, held tenaciously to his oil-drop value for e for a long while, claiming that the x-ray method only sampled the crystal lattice spacing at the *surface* of a crystal and that perhaps this very special location might differ from the rest. I became involved a little later, by doing a series of experiments with V. L. Bollman and others, which blocked this last possible avenue of escape.[5] It was Professor Birge who called our chief's attention to my paper and I have never forgotten the gentlemanly way in which Dr. Millikan forthwith called me to his office specifically to acknowledge that my reasoning was incontrovertible and that his value of e must require revision. That was in 1936.

Jesse DuMond, professor emeritus at California Institute of Technology, started pursuing the fundamental constants about a decade after he came to Caltech as a teaching fellow in 1921. In a letter about this paper, which is based on a talk presented last April to the American Physical Society in Washington, he wrote, "The ideas are important, and I sense at present a lazy trend away from them. Some important things in the history of physics have turned on minute discrepancies or effects."

In 1939 the simplifying concept of what I would like to call "constants space" occurred to me.[6] Suppose we select some set of atomic constants, n in number, which we think of as the unknown quantities we wish to determine from experimental data. Most often, said data do not give these unknowns themselves directly. More frequently the data give functions of several of the unknowns simultaneously. Let us visualize a rectangular coordinate system in n-space, the space of the unknowns we have selected. For example, we might plot in a 3-space the three constants, e, m, and h, so that each point corresponds to a set of values of those three quantities. The selection of the particular quantities, e, m, and h, for the principal axes, is arbitrary, not uniquely indicated. Other axes in constants space could equally well be chosen, which would correspond to other constants functionally related to e, m, and h, for example α, the fine-structure constant, e, the fundamental charge, and N, the Avogadro number. Now the quantities measured experimentally, which are functions of our primary variables will appear as curved surfaces in this space. Three such curved surfaces intersecting in the 3-space may suffice to determine a common point. Usually, however, we have more experimental data than just sufficient, and the plurality of curved surfaces will intersect somewhat ambiguously in a small region of the space in which, except for degenerate cases, any set of three surfaces may define a slightly different point.

Let us simplify the picture a little more. I first remind you that we already know pretty well from previous work the approximate location of our point, say, e, m, and h. In 1929 the uncertainties as to the correct values of such constants were of the order of a fraction of a percent; today they are of the order of, at most, tens of parts per million. Now, on such a small scale of variation as this, the curvature in our curve surfaces, all of which are smooth, well-behaved functions, is entirely negligible, and for all practical purposes we can replace the surfaces by their tangent planes in the region of interest. What we do then, in practice, is to select a new origin of coordinates at a point e_o, m_o, h_o, values which we know differ little from the true values we are seeking, and we then re-express our entire situation in terms of new variables, x_e, x_m, x_h, which are the *relative deviations* from these origin values expressed, for example, in parts per million. Essentially what we have done is to replot our situation in logarithmic coordinates.

It is important to realize that we may have degenerate cases in which a set of three planes does not determine a unique point because they may either lie parallel to each other without intersecting at all, or the three pairs may intersect in three parallel lines. This last case, which occurs

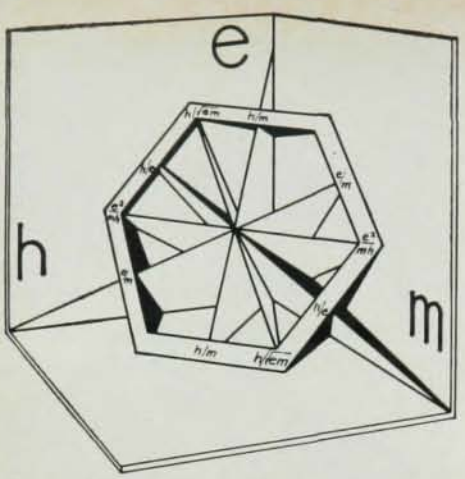
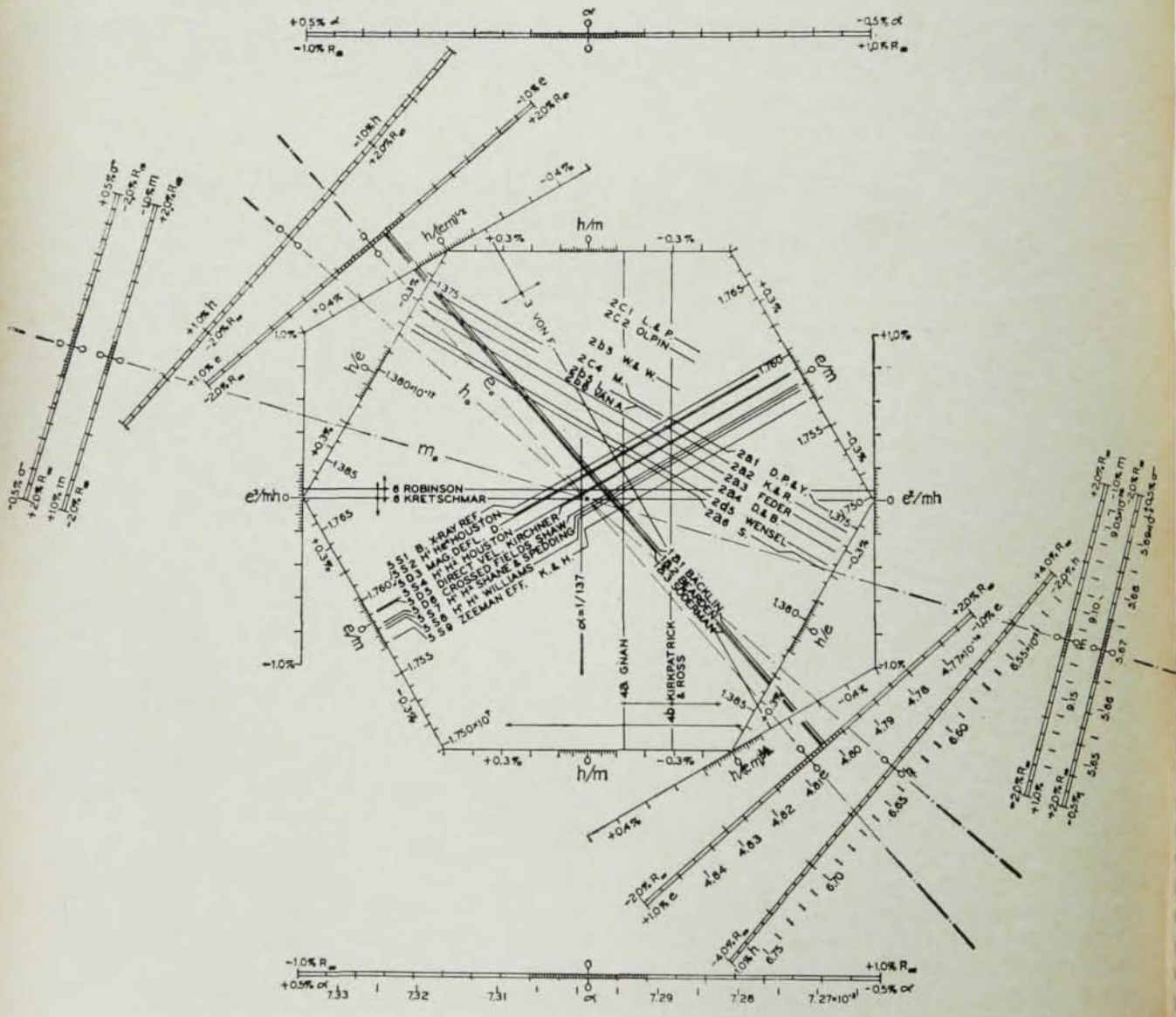


Fig. 1. Five cozonal planes in constants space lie parallel to the space diagonal of e, m, h axes. For simplicity they are shown coinciding with that common axis.

quite often, I have called the cozonal case. A set of five cozonal planes for the functions, $e^2/(mh)$, h/e, $h/(em)^{\frac{1}{2}}$, h/m, and e/m is shown in Fig. 1.

If the results of different experiments are represented by planes in constants space, and the planes fail to determine the same single common point in the space, we say the experiments are not consistent. Each plane, however, representing as it does the numerical result of a physical measurement, has associated with it a standard deviation, that is, an estimated root-mean-square deviation from the mean, indicative of the accuracy estimate of the numerical result. This can be visualized as a thickness, normal to the plane, indicative of its indeterminacy of position. We say then, that a determination is significantly inconsistent with a point in our space if the plane misses the point by a large distance relative to its own standard deviation.

Fig. 2. Consistency diagram for 1940 data. Individual data are to be found in Ref. 7.



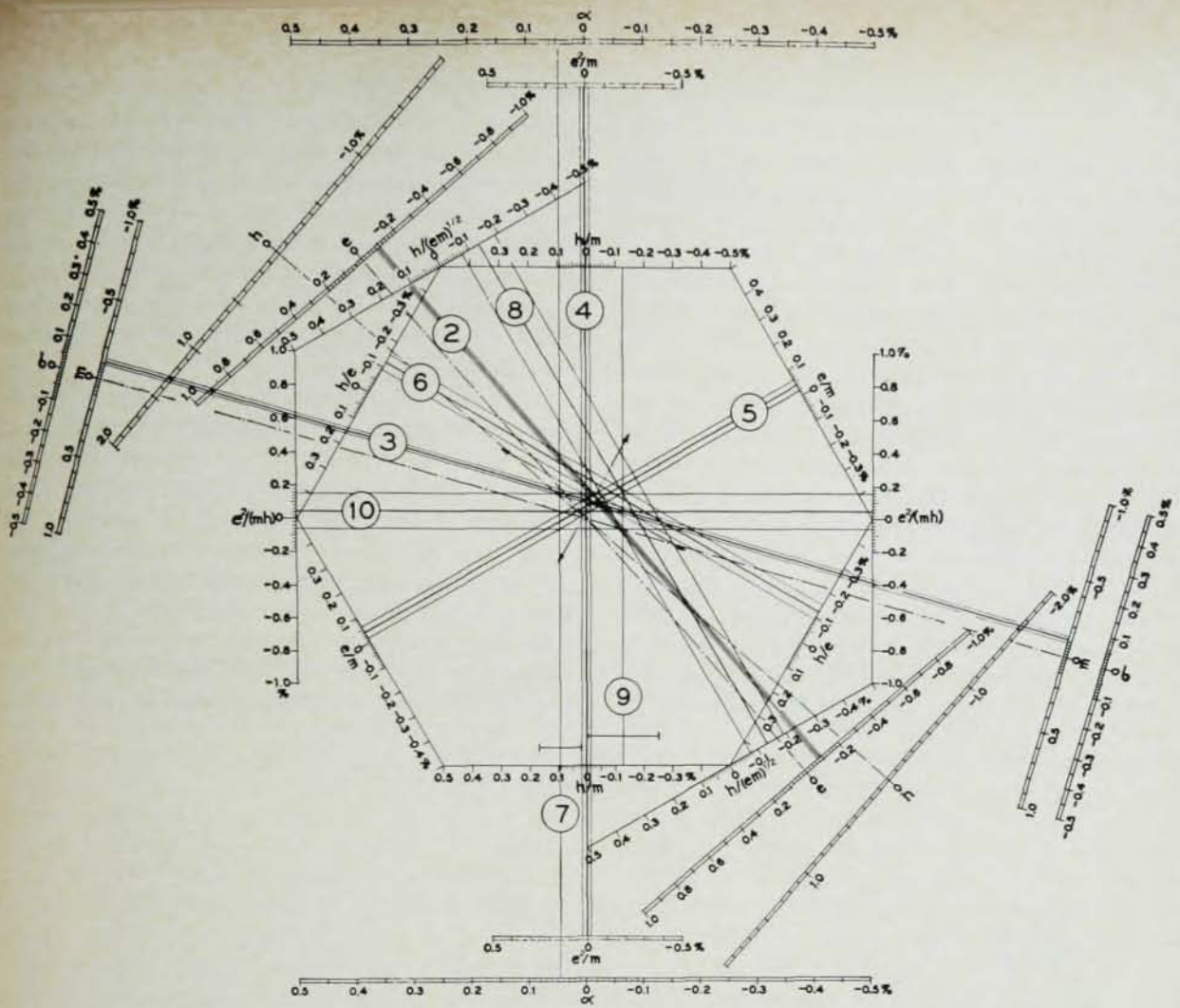


Fig. 3. Consistency diagram for 1947 data, which are in Ref. 8

Clearly we can examine two-dimensional cross sections through our variously tilted planes in this constants space which I have just described, and if we choose these cross sections judiciously, they can be very revealing as regards the state of consistency or inconsistency of our knowledge of the constants at any given epoch.

Figure 1 shows the orientation of some of the planes for one of these early consistency diagrams, which I prepared in 1940.[7] At that time the experimental measurements which had been performed happened to yield a large number of planes in constants space which were cozonal, that is to say, parallel to a common line—the space diagonal, in fact, of the rectangular coordinates e, m, h. Fortunately, there existed one measurement, the Rydberg constant, $R_\infty = 2\pi^2 me^4/(h^3 c)$, which furnished us with a plane intersecting the space-diagonal direction, and this removed the degeneracy. Figure 2 shows the resulting consistency diagram in which the cross section is taken normal to the cozonal axis (the above-mentioned space diagonal). The whole length of each side of the hexagon corresponds to a range of variation of plus or minus one-half of one percent. Figure 2 is a prime example of obviously highly inconsistent data. The input data are too numerous to permit showing the standard deviation ranges of each datum without confusion. To every intersection of any pair of lines on this graph there corresponds a possible set of values for the three quantities, e, m, and h, and from each of these sets almost all of the fundamental constants and conversion factors of physics and chemistry can be computed. I made no attempt to solve for least squares adjusted "best" values from these 1940 data but contented myself with merely illustrating in this way the chaotic state of inconsistency which then prevailed. Figure 3 shows a similar consistency graph based on our selection of available data in 1947.[8] Here the standard deviation ranges are shown by lines on either side of the central value line of

each determination. The consistency here was notably better and Dr. Cohen and I made a least-squares adjustment which resulted in the tiny ellipse of error visible near the center of the graph. The scale here is the same as in the previous figure, ± 0.5 percent for the sides of the hexagon.

Some three years later a huge improvement in the accuracy of available input data had occurred. Such experiments as Gardner and Purcell's cyclotron frequency of the electron, the "omegatron" determination of the cyclotron frequency of the proton by Hipple, Sommer, and Thomas, and of the gyromagnetic ratio of the proton by Thomas, Driscoll, and Hipple had been performed. These, together with x-ray determinations by J. A. Bearden and co-workers of the short wavelength limit of the continuous x-ray spectrum, and by Bearden, Bäcklin, Söderman, Tyrén, and others of the Avogadro number by the x-ray crystal-density method, furnished the input data for the consistency chart[9] shown in Fig. 4. This depicts the state of consistency in 1950. Here the linear scale of the diagram is multiplied one-hundred fold relative to its predecessor so that the length of each side of the hexagon corresponds to a deviation of plus or minus five parts in ten thousand. The tiny ellipse of error, which was barely visible on the preceding 1947 chart, is shown here to correct scale, in position and orientation, for comparison with the new, much smaller, ellipse of error for this 1950 situation.

The last three figures illustrate just three of the earlier stages of the pilgrimage, first alone, later with my able friend, E. Richard Cohen, in search of better and more consistent values of the constants. After 1950 we found it necessary for some time to include as many as three or four unknowns in our least-squares adjustments. Dr. Cohen has, I believe, just completed a new one with five unknowns. In such situations we are obliged to abandon the use of two-dimensional graphs, save as illustrations, and to rely much

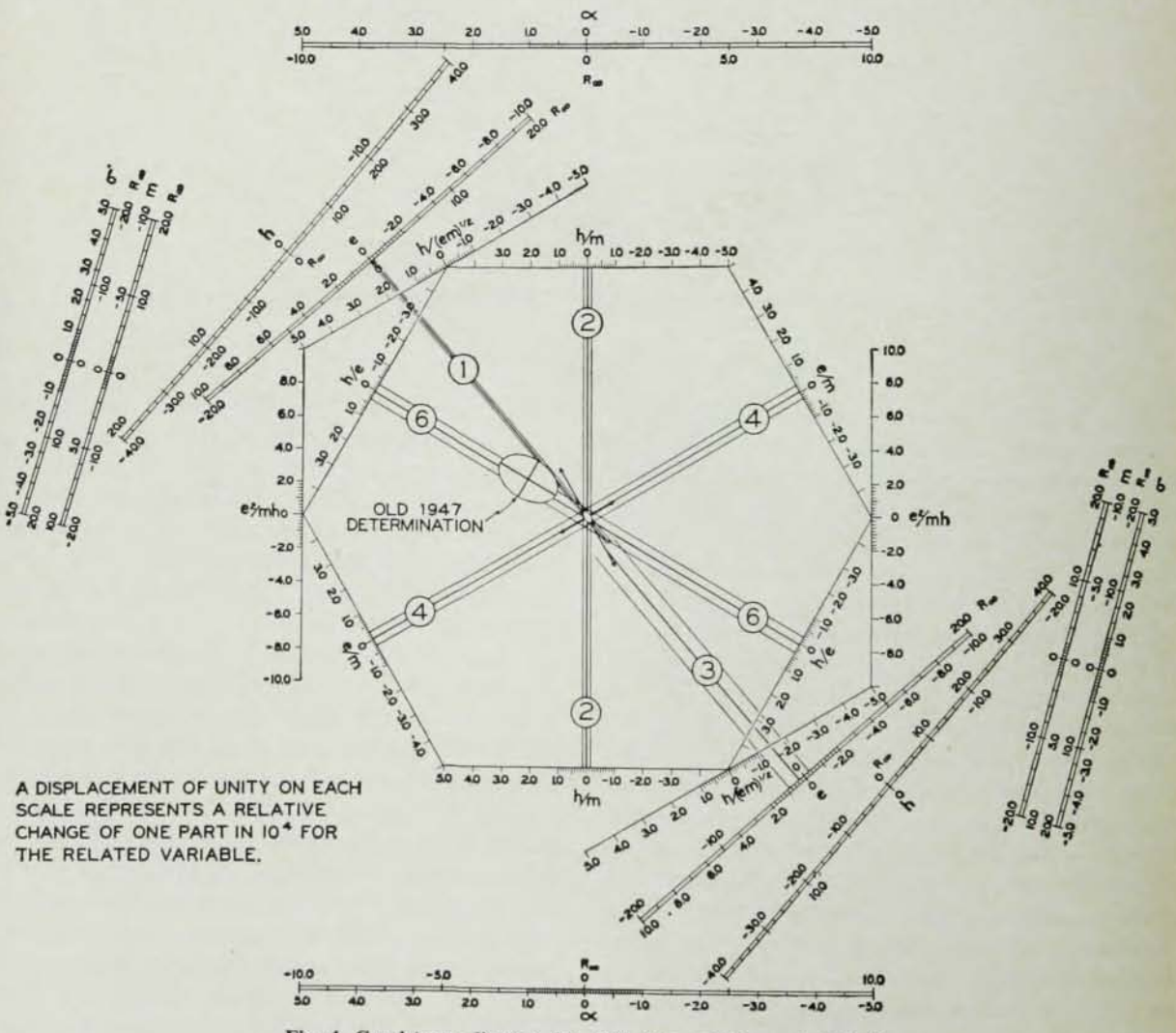


Fig. 4. Consistency diagram for 1950 data, which are in Ref. 9

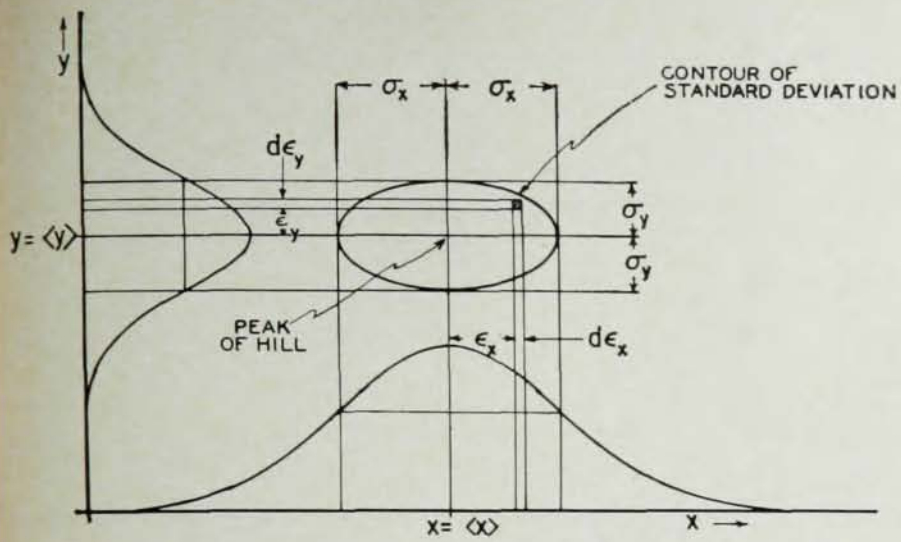


Fig. 5. Error hill for two uncorrelated stochastic variables.

more on analytical methods of examining consistency. Fortunately, the use of electronic digital computers has saved us much drudgery of numerical computation. A least-squares adjustment which might take upward of a week by hand computation with a desk computer, can now be done in a second or so, once programmed, and it is possible to do several hundred of them, using different combinations of input data, to test which subset combinations yield the best consistency. Such an exploration of different combinations of available data we have come to call an analysis of variance. Our first analysis of variance was made in 1955.[10]

We started our latest adjustment of the constants in 1960 at the suggestion of the National Research Council Committee on Constants, whose chairman is A. G. McNish, now chief of the Division of Metrology of the National Bureau of Standards. A demand had arisen at about that time for a new set of values of the constants to be recommended for general use. This demand was triggered, as it were, by two new conventional changes: (1) the adoption in 1955 of a new slightly different scale of temperature, the Celsius scale, and (2) the adoption in 1961 of the unified scale of atomic weights. A far more interesting reason, from our point of view, for a complete reanalysis of the constants was that since 1955, the date of our last previous effort, a wealth of new and much improved experimental data had become available and errors in some of the old input data of 1955 had been discovered; they were mostly errors of experiment, as usual, but one important case was an error of theory, concerning the electron magnetic moment anomaly ratio, and requiring revision of the coefficient in the second-order term of the theoretically derived expansion of that ratio in powers of α.[11]

About 1950 Dr. Cohen and I had become keenly aware of the importance of error-statistical correlation in working with weighted averages and least-squares adjustments. The procedure of forming a least-squares adjustment is, in fact, nothing more than the formation of a weighted average in more dimensions than one. It is an analytical procedure for finding the centroid point in multidimensional space.

We all know that the numerical results of a physically measured quantity, say, x, are subject to errors of observation and that the uncertainties in such numbers, x, can be idealized in mathematical statistics as distributions, Gaussian or otherwise, about a mean value, $<x>$, having a root-mean-square measure of spread, the so-called standard deviation, σ_x. I would like to adopt the word stochastic to describe this idealized concept of such quantities, because if I call them erroneous quantities or imprecise quantities this may be taken to mean that they are systematically erroneous, that the entire error distribution is biased or shifted to a significantly incorrect position. The word stochastic implies a mathematical idealization or fiction: the idea that there exists a "parent distribution" such that, if we made enough observations of the stochastic quantity, the distribution of the values obtained would approach this parent distribution curve.

Now, suppose we make completely independent sets of observations on two different stochastic quantities, x and y. Figure 5 shows the two distribution curves with standard deviations, σ_x and σ_y, as measures of their root-mean-square spread. The ellipse is the contour of standard deviation

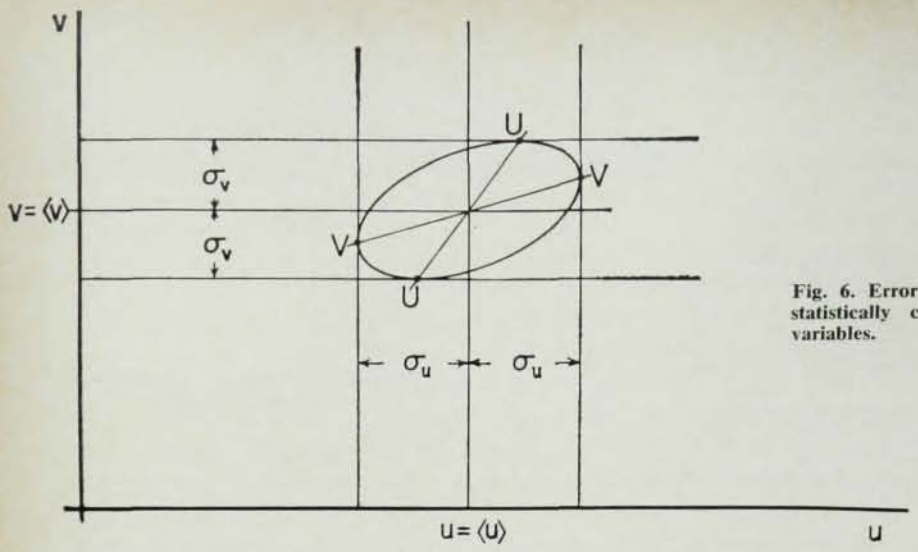


Fig. 6. Error hill for two error-statistically correlated stochastic variables.

of the error-hill, a two-dimensional distribution exhibiting, by its height normal to the plane of this figure, the probability that a deviation ϵ_x and simultaneously a deviation ϵ_y will be observed. The fact that the principal axes of the ellipse are parallel to the axes of the two stochastic quantities, x and y, is evidence that the two error distributions are statistically independent or uncorrelated. The peak position of a cross section through the hill taken at an arbitrary but constant value of x does not shift in the y-direction as we vary the value of x.

But now, suppose we multiply both stochastics, x and y, by a third stochastic, z, whose error distribution has a standard deviation, σ_z. Immediately the two products, $u = xz$ and $v = yz$, now become correlated stochastic quantities. Figure 6 illustrates this situation. The error hill now has elliptical contours with oblique principal axes. The lines UU and VV are called regression lines; VV, for example, is the locus of the peak positions of cross sections taken through the hill parallel to v as we vary the u position of the cross section. The ratio of the displacement of the peak at V on the standard-deviation contour to the corresponding standards deviation, σ_v, is called the correlation coefficient, r_{uv}. As shown in this figure, it is positive and slightly less than one-half. It always has the same value for v relative to u as for u relative to v.

Now the important facts that we first became aware of in 1950 are (1) that the ordinary textbook procedure for effecting a least-squares adjustment is in need of considerable modification and generalization and becomes much more complicated if the equations of observation are statistically nonindependent (that is to say, correlated), and (2) the same may be said of the familiar formulas for computing propagated errors of functions of correlated stochastics: the generalized formula of error propagation must be used. It contains additional terms which take into account the correlations between all pairs of stochastic quantities. Surprisingly few experimental physicists seem to be aware of this pitfall.

Figure 7 shows the generalized formula of error propagation,[12] which, as you see, contains cross-product terms in addition to those involving squares of the standard deviations of the different

$$y = f(x_1, x_2, \cdots x_n)$$

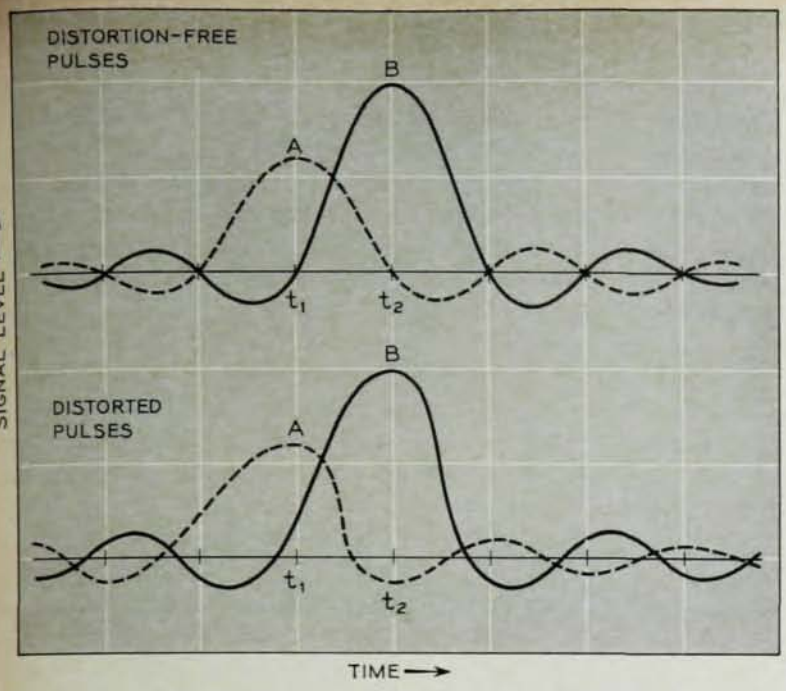
VARIANCE MATRIX

$$\begin{array}{c} \\ x_1 \\ x_2 \\ \vdots \\ x_n \end{array} \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & \cdots & x_n \end{array} \\ \begin{bmatrix} v_{11}=\sigma_1^2 & v_{21} & \cdots & v_{n1} \\ v_{12} & v_{22}=\sigma_2^2 & \cdots & v_{n2} \\ \vdots & \vdots & & \vdots \\ v_{1n} & v_{2n} & \cdots & v_{nn}=\sigma_n^2 \end{bmatrix} \end{array}$$

$$\sigma_y^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial y}{\partial x_i} v_{ij} \frac{\partial y}{\partial x_j}$$

$$\sigma_y^2 = \sum_{i=1}^{n} \left(\frac{\partial y}{\partial x_i}\right)^2 \sigma_i^2 + 2\sum_{i<j} v_{ij} \frac{\partial y}{\partial x_i} \frac{\partial y}{\partial x_j}$$

Fig. 7. Generalized error-propagation formulas for the variance of a function of correlated stochastics that have the variance matrix shown.



Report from

BELL LABORATORIES

Concept of the data-distortion problem in simplified form: Two data pulses at top are undistorted because their "tails" all pass through zero signal level at sampling times t_1, t_2, etc. Thus, receiving circuit "looking" at the signal at time t_2 would "see" signal energy from pulse B only. Distorted pulses at bottom, however, have tails which do not pass through zero at sampling times. In this case, receiver at time t_2 will see energy from both pulse A and pulse B and might register a false signal level. Such distortion must be reduced to a minimum to achieve high data transmission rates.

AUTOMATIC EQUALIZER MINIMIZES DATA DISTORTION

A communication signal arriving at its destination is never a perfect replica of the original. There is always some distortion, and if this distortion exceeds acceptable limits, it must be reduced by a process known as equalization.

Equalization increases the rate at which data pulses can be transmitted. Ideally, the equalization should also adapt rapidly to changing transmission characteristics, which are caused by varying temperature, humidity and other factors. Otherwise, distortion may cause receiving circuits to register false values for the data pulses (see above drawing).

To solve this problem a new data equalizer promising increased data rates—up to a threefold increase on voice-telephone channels—has been devised at Bell Telephone Laboratories. With this new equalizer, test pulses cause a series of adjustments to be made in the settings of equalizer attenuators. These adjustments, impossible to perform rapidly by hand, are performed automatically by control circuitry. As a result, the equalizer quickly reaches a condition of minimum data distortion. Later, when the transmission characteristics of the line change, the equalizer automatically adapts to the changes by making additional adjustments that keep the attenuators at their optimum settings.

Experimental automatic data equalizer devised at Bell Laboratories. Control section consists of the circuit packages; the package being inspected is one of 12 attenuators, the settings of which determine the degree of distortion-correction of the equalizer.

Bell Telephone Laboratories
Research and Development Unit of the Bell System

"Steepest descent" minimization

The new data equalizer was made possible by a discovery by R. W. Lucky at Bell Laboratories that a technique of minimizing mathematical functions is applicable to the problem of data equalization. Known as the "steepest descent" technique, it is analogous to a hiker desiring to climb down a hill in minimum time. In the equalizer application, it was shown that the steepest descent technique results in the true minimum, and not a local or relative minimum of the function. It was also shown that an equalizer based on this technique could be built with simple control circuitry. An experimental model of the equalizer (see photo) uses a 12-tap delay line in a transversal filter with an adjustable attenuator at each tap. The control circuitry extracts information from each test pulse, and for each pulse adjusts all attenuators by small steps, each step calculated to reduce distortion in the direction toward the minimum.

error-contributing stochastics. These additional terms may be either positive or negative and their omission can seriously falsify calculation of propagated errors resulting in either too large or too small an error estimate.

Another important point to realize is that the output values of a least-squares adjustment are necessarily in general correlated stochastics. Therefore, it is never sufficient to give only the standard deviations of such output values. If the values are to be used in further formulas to compute functions dependent on several of them, one needs the whole error matrix, the inverse of the coefficient matrix of the normal equations, including all the off-diagonal terms.

I purposely started by showing you how two completely independently observed stochastics, x and y, can readily be converted into two correlated stochastics, u and v, simply by multiplying x and y by a common factor, z, which is also stochastic. Such a situation is likely to arise in the work of analyzing the constants because of the unavoidable presence in physics of certain stochastic conversion factors which cannot be eliminated. Conversion factors are a nuisance and nobody likes them, but in the present state of the arts of high-precision metrology, some of these conversion coefficients must be taken into account explicitly as separate operational steps, if we wish to push precision to its utmost, precisely because they introduce correlations between otherwise error-statistically independent quantities. This ineluctibility of certain conversion factors is, I fear, not as generally realized as it should be. Table 1 lists some of these conversion factors which cannot at present be safely eliminated, though of course improvements in metrological technique may later change this situation. Such an indispensable conversion factor is required whenever metrological techniques within each of two different fields of physics (such for example, as length measure and liquid volume measure) have attained an accuracy permitting intercomparison of measured quantities within either given field with more precision than the precision with which the conversion factor connecting the two fields has been determined.

Table 1. Presently inescapable conversion factors (as of 1965)

1. Linear-to-volumetric (liquid) measure
2. Relative-to-absolute nuclidic mass scales (Avogadro number N)
3. X-unit-to-milliangstrom-unit conversion factor (Λ)
4. Mass-to-force units (absolute g at place where force is weighed)
5. "As-maintained"-to-unqualified absolute electrical units

I shall not discuss all the examples of this situation given in the table, but one of them is of some importance to the present status of the fundamental constants. This is the conversion constant between the fundamental length units, in terms of the angstrom or meter, on the one hand, and the arbitrary scale of length units in terms of which x-ray wavelengths are measured. This conversion factor, which I shall call Λ, is defined such that the wavelength of an x-ray emission line, expressed on the nominal x-ray wavelength (or so-called x-unit) scale must be multiplied by Λ to convert it to the corresponding wavelength in milliangstrom units. Recently this conversion constant, Λ, has been redetermined by some five different methods, some of them direct, others indirect, the best probably being the work of Henins and Bearden with silicon crystals,[13] but even now the claimed relative standard deviation of these Λ measurements is still at least five or more times greater than the best attainable in measurements of x-ray emission-line wavelengths relative to each other by means of crystal diffraction.

The x-unit was intended by Manne Siegbahn, who proposed its use, to represent approximately 10^{-11} centimeters, or a milliangstrom unit. X-ray emission-line wavelengths, as everyone knows, are most accurately measured by means of x-ray diffraction from the lattice planes of natural or artificial crystals, so that in an operational sense the meter stick for comparing the ratio of any two emission line wavelengths is the atomic interplanar spacing in the crystal. If we always use the same crystal, under the same conditions, this furnishes a very reliable way of intercomparing the different wavelengths of hundreds of lines. Using crystals of different species for different wavelength ranges, the ranges being chosen so as to overlap each other, one can thus tabulate, on a nominal but highly accurate relative scale, the wavelengths of perhaps 3000 x-ray emission lines over a very broad range with relative accuracies approaching perhaps ± 1 part per million, or even better.

Now the great pioneer in precision x-ray spectroscopy, Manne Siegbahn, undoubtedly realized that he could not compute in absolute length units, centimeter or milliangstrom, the interplanar spacing of the cleavage planes of a calcite crystal from the then ill-known values of its crystal density, molecular weight of $CaCO_3$, the geometry of the unit cell, and the Avogadro number (that was

proven performance... with FIELDIAL™

If you're not already using a FIELDIAL* Magnetic Field Regulator, ask one of your friends about its performance and convenience. From their years of experience, hundreds of scientists throughout the world will attest to the advantages of the FIELDIAL Regulator's full-time field control with continuous protection against external a-c and d-c field changes. The result: volumes of more accurate and meaningful data.

And never underestimate the advantages of the FIELDIAL Regulator's solid-state, field-sensing probe. Small, compact, and weighing virtually nothing, it attaches quickly and easily to any magnet pole cap. You can forget about it—it's out of your way. You simply dial the desired field and start your experiment.

Ask one of your colleagues for a demonstration or write us for complete descriptive information and performance specifications.

*Trademark of Varian Associates

VARIAN ASSOCIATES ANALYTICAL INSTRUMENT DIVISION
PALO ALTO, CALIF. • ZUG, SWITZERLAND

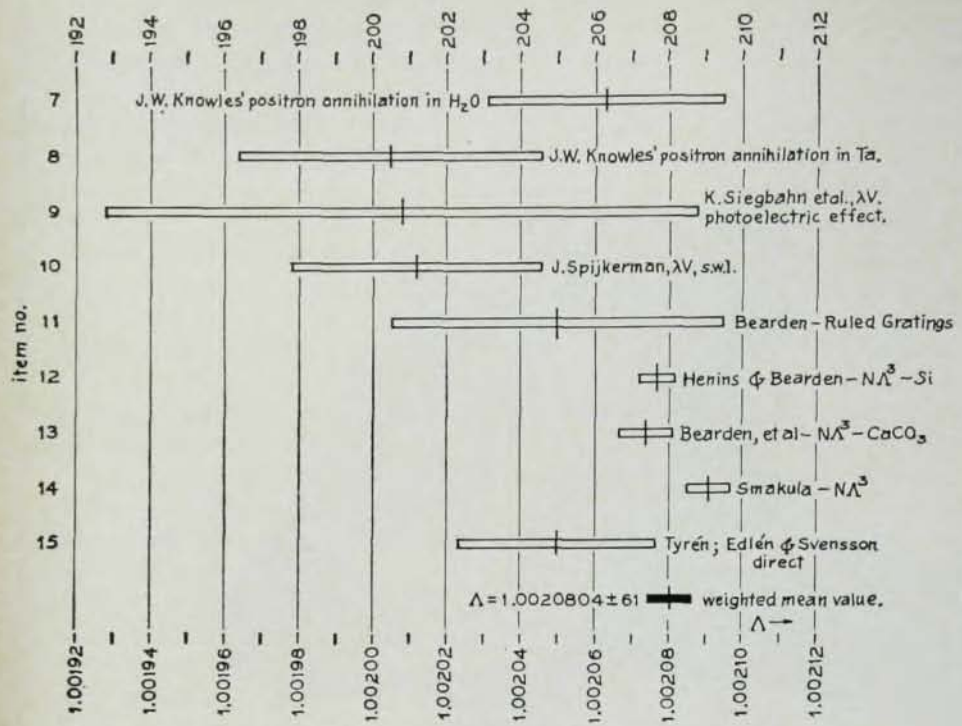


Fig. 8. Diverse results on the x-unit-to-milliangstrom conversion constant. In-text references show sources of data. In calculating these results 1963 recommended values of other constants have been used.

circa 1930), by any means as accurately and surely as he could intercompare x-ray wavelengths using a specific calcite interplanar spacing as a meter stick. Accordingly, he decided to use his arbitrary meter stick, the crystal species calcite, as a definition of the x-unit*. Later, precision research[14] by J. A. Bearden, M. E. Straumanis, and many others, has shown that this was unfortunate since perfectly clear, beautiful-looking calcite can vary in grating spacing from sample to sample by as much as 100 ppm and, even among chemically tested purest specimens, by tens of parts per million. The danger is clear that early workers using different calcite crystals, unaware of this variability, have reported x-ray emission-line wavelengths which are not really expressed in the same units but which may, nevertheless, have found their way into the wavelength tables. Thus it was a serious mistake to define the arbitrary unit in terms of a crystal species. Instead, the unit should have been defined by tying it to a judiciously chosen standard x-ray emission line since, with proper precautions, these are much more reproducible. I am happy to see that now, for the first time in history, this preferable method which I have been advocating for the last eight years, has been adopted by Dr. Bearden and his co-workers in their newly issued tables of x-ray wavelengths.[15] Therein they define the x-unit by assigning the numerical value, 208.5770 x.u. exactly, to the peak intensity wavelength of the W K α_1 x-ray emission line. They also define another unit, Å, similarly tied to the tungsten line, by the statement that the peak intensity wavelength of the W K α_1 line is 0.209010 Å. The numeric here has been selected by Professor Bearden and his group as a value which they believe, on the basis of their and others' measurements of the conversion factors, Λ, to be close to one angstrom (their assigned probable error is ±5 ppm). However, it is important to realize that they define their Å unit operationally in terms of the peak wavelength of the tungsten emission line, an operation repeatable probably to better than ±1 ppm, whereas the conversion factor from Å to angstrom, it must be emphasized, is still a stochastic quantity known to no better relative accuracy than the conversion factor, Λ, from x-units to milliangstrom units, since, in fact, both repose on the same experimental determinations. The graph in Fig. 8 illustrates the great diversity of sources of recently obtained direct and indirect information on this conversion constant, Λ. Reading from top down we have items 7 and 8, work on the annihilation radiation by J. W. Knowles;[16] item 9, work using Kai Siegbahn's magnetic β-ray spectroscopic method to determine energy differences between x-ray lines by direct retardation or acceleration of photo-ejected electrons;[17]

* He clearly stated the convention (in the 1931 edition of his text) that he would take the effective grating constant for first order reflection of the cleavage planes of calcite at 18° C to be $d_{18}° = 3029.04$ x-units.

item 10, a recent determination by Spijkerman of the quantum limit of the continuous x-ray spectrum from a mercury vapor target;[18] item 11, direct information on Λ from ruled-grating diffraction of x-ray emission lines by J. A. Bearden;[1] items 12, 13, and 14, sundry x-ray crystal-density determinations [13,19,20] of $N\Lambda^3$; and finally item 15, a value of Λ implied by the recent work of Edlén and Svensson[21] who have, as a result of our urging since 1956,[22] remeasured Tyrén's 1940 x-ray and spark spectra[4] and corrected his results for the Lamb shift. The scale of Λ on this graph covers a variation of 200 parts per million, from 1.00198 to 1.00218, and the various values range over 85 parts per million.

Out of the nine items numbered 7 to 15, only two, 11 and 15, however, are direct determinations of the conversion constant, Λ, by means of ruled-grating diffraction. All the others depend, not only on the observer's measurements, but also on other constants, such as N, the Avogadro number, or α, the fine structure constant. Note that this statement applies to items 12, 13, and 14, by far the most influential contributors to the weighted mean. The 1963 adjusted values of N and α were used in calculating these results and these quantities are intimately interlocked and subject to possible modifications, especially because of an uncertainty regarding α, which I shall discuss further presently. For every part per million relative increase in α, the three important values of Λ, items numbered 12, 13, and 14, will have to be increased by the same number of parts per million, while the values numbered 7, 8, 9, and 10 will increase twice as many parts per million. Clearly this will improve slightly the consistency of the results. The weighted mean value will show about the same relative increase as does α. If such changes are large, they would involve either revision of Dr. Bearden's definition of his Å and reprinting of his wavelength tables, or application of a revised conversion factor to correct the wavelengths when highest precision is required.

When we started in 1960 to work on what finally became the 1963 least-squares adjustment, Dr. Cohen proposed, with my hearty approval, to try to be even more critically rigorous than we had been in the past in the matter of detection and elimination of systematically erroneous input data. Suppose we have two independently measured results, purporting to be measurements of the same constant, to each of which a standard deviation error estimate has been attached by the authors, but which are so significantly discrepant as to be distinctly resolved, that is to say, they differ by an amount large in comparison to the expected standard deviation of the difference. Figure 9 shows how the two error distributions might look. I agree with Dr. Cohen in asserting that it

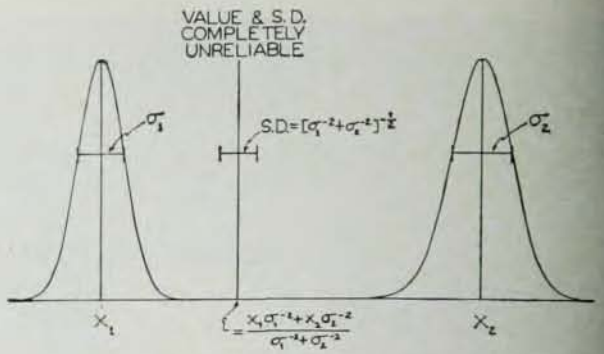
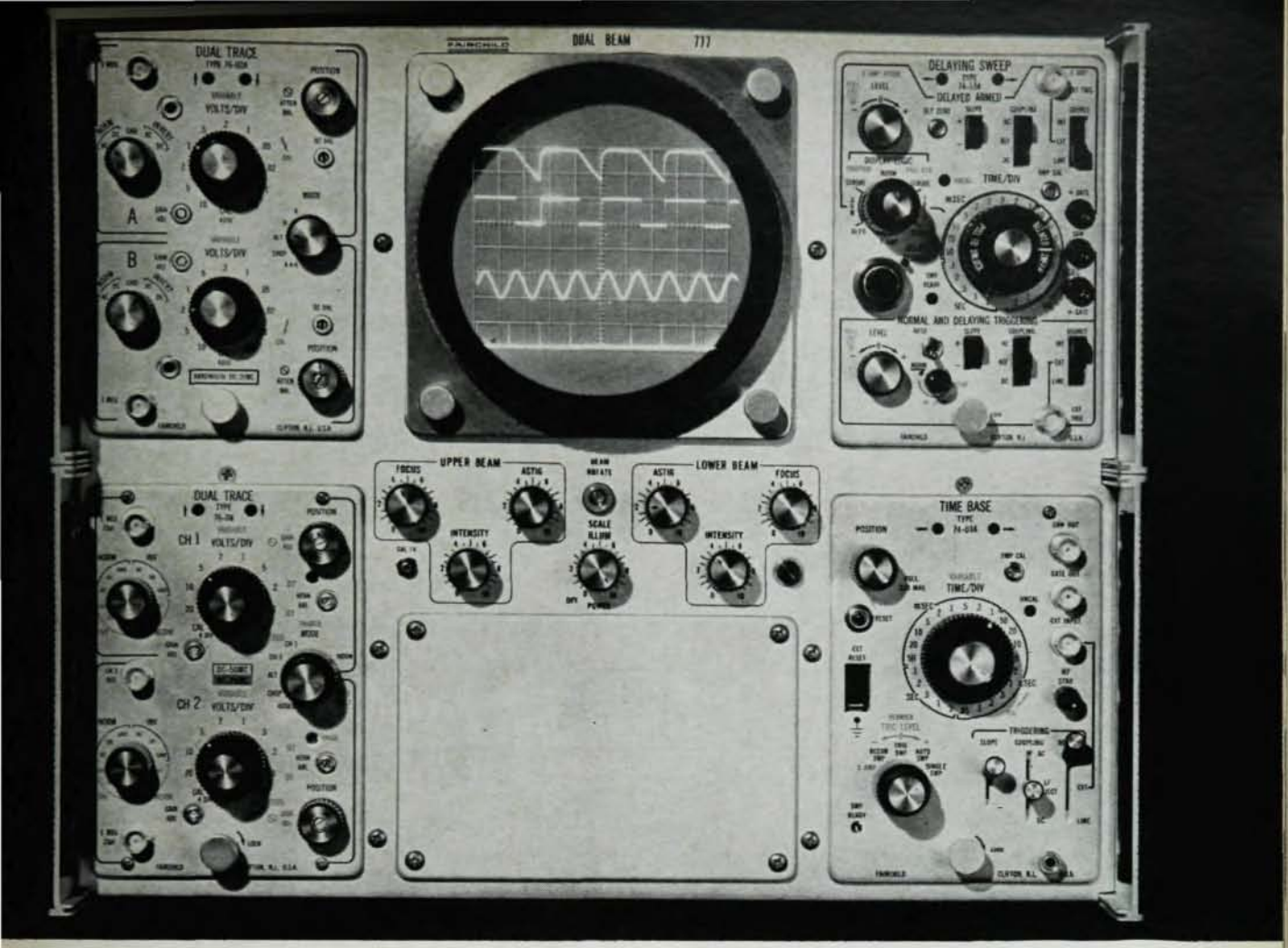


Fig. 9. The centroid of two clearly discrepant stochastics purporting to be measurements of the same quantity is unreliable.

is incorrect (in fact, a travesty on the principle of least squares) to attach meaningful significance to a weighted mean of these two values with weights inversely as the squares of the two standard deviations. The patently significant inconsistency of these two stochastics guarantees that at least one of the four quantities, x_1, x_2, σ_1, or σ_2, must be wrong. The weighted average employing these quantities must therefore be unreliable. We have only two choices: either to discard both data, or to re-study the details of both experiments with great care, correspond with the authors, if necessary visit their laboratories, and try to think of every possible source of error which might have been overlooked, so as to choose between them.

For our 1963 adjustment, Dr. Cohen devised some very ingenious tests.[23] These are based on the theory of least squares and are designed to detect discrepant items of input data by a process analogous to locating and measuring the stored energy of internal strain in the members of an overdetermined mechanical structure. I must emphasize that we in no case, however, discarded any eligible input datum on the sole grounds of its being outlying from the group. We invariably took great pains to find a valid reason in the experiment itself which might also suffice to explain why it might be likely to be in error.

For the 1963 adjustment, Dr. Cohen and I devoted much effort to collecting and recalculating all the x-ray data we could find available at that time bearing on the combination $N\Lambda^3$, the product of the Avogadro number, N, into the cube of the



HOW MANY SCOPES CAN THIS ONE REPLACE?

A sizable number, depending upon the range of applications. For this is the Fairchild 777—the most versatile of all industrial scopes. The 777 is a dual beam, dual trace scope in which any four of 22 plug-ins are completely interchangeable in both X and Y cavities. These same plug-ins fit all Fairchild 765H Series scopes. They include DC-100 mc bandwidth, spectrum analyzer and raster display capabilities, sensitivity to 500 μv/cm, risetime to 3.5 ns.

Other features of the 777 include 6 x 10 cm display area for each beam with 5 cm overlap between beams for optimum resolution... unique 13 kv CRT with four independent deflection structures... solid state circuitry (with all deflection circuitry in the plug-ins)... light weight (44 lbs.)... environmentalized for rugged applications. Price (main frame): $1,600 f.o.b. Clifton, N.J.

The 777 illustrates the Fairchild concept of value through versatility. One scope doing many tasks is only part of it. Future state-of-the-art capability is equally important because it helps you curb the high cost of Technological Obsolescence. And finally, service. Fairchild has more service centers than any other scope manufacturer.

Ask your Fairchild Field Engineer for details on this and other new generation Fairchild scopes. Or write to Fairchild Instrumentation, 750 Bloomfield Ave., Clifton, N.J.

BEAT T.O. WITH FAIRCHILD

*Technological Obsolescence

FAIRCHILD

INSTRUMENTATION

A DIVISION OF FAIRCHILD CAMERA AND INSTRUMENT CORPORATION

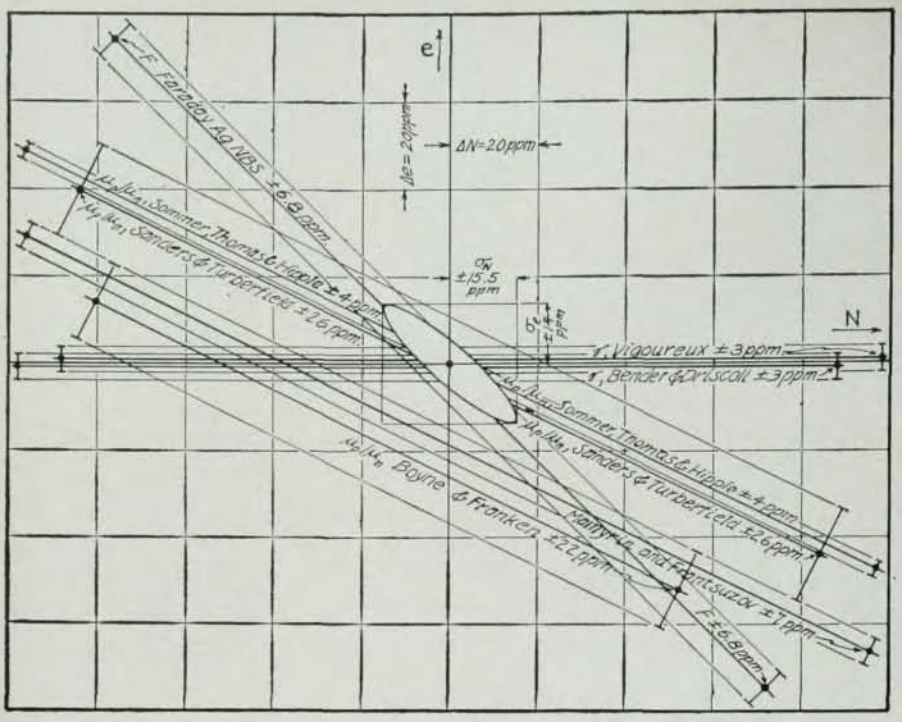


Fig. 10. Consistency graph for input data of 1963 adjustment. Central origin corresponds to e = 4.80298 × 10^{-10} esu, N = 6.02252 × 10^{23} per mole, and a scale of ^{12}C = 12. This is a cross section in the plane on which alpha is 7.29720 × 10^{-3}.

x-unit-to-milliangstrom unit conversion factor, Å. NA^3 is the universally constant quantity which should result, when the molecular weight of the unit cell of an ideally perfect crystal is divided by the product of the macroscopically measured crystal density into the volume of the unit cell, with the latter volume measured in cubic x-units, as determined from x-ray diffraction work. Two internally concordant sets of eight separate results, each on NA^3, one set made with the Mo K α_1 line, the other with the Cu K α_1 line, failed to agree by a very unpalatable discrepancy and, since the NRC committee was pressing us in 1962 to give them a set of adjusted values promptly, Dr. Cohen and I decided that the only safe course was to dump all of these discrepant x-ray data overboard, since no clear criteria were available to permit us to save any part and discard the rest.[24,25]

The price we paid for rejecting the ambiguous x-ray data was that we were left with a set of degenerate observational equations which fell into two disjoint groups. One group consisted of five very satisfactorily interconsistent input data. These, however, constituted a cozonal set determining a line in the constants space, but not a point. The other group consisted of two discrepant sources of information on the fine-structure constant α, namely the measurement by Dayhoff, Triebwasser, and Lamb of the fine-structure splitting in deuterium[26] and the measurement by Ramsey and associates of the hyperfine splitting in hydrogen with the hydrogen maser.[27] The latter resulted in a value of α 26 ppm larger than the former. These two values determined two different parallel planes intersecting the line determined by the cozonal group in two different points. This situation afforded no consistency criterion whatever to indicate which α was preferable. We chose to reject the hyperfine-structure α because of an uncertain, theoretically calculated, correction term (for the internal field structure of the proton) which is needed in its calculation, but which does not enter in the fine-structure case.[28] Figure 10 shows a cross section of the constants space in the Lamb-value plane, the one which we adopted. The axes corresponding to the fundamental charge, e, and the Avogadro number, N, in this plane are shown and the squares correspond to changes in these two quantities of 20 ppm. The ellipse of error is seen to be much elongated. The cozonal planes intersect the α plane obliquely, giving lines of three different slopes. The horizontal lines correspond to determinations by Bender and Driscoll[29] and by Vigoureux[30] of γ, the gyromagnetic ratio of the proton. The less steeply sloping lines are measurements of the magnetic moment of the proton in nuclear magnetons by Hipple, Sommer, and Thomas[31] and by Sanders and Turberfield.[32] Two additional data in this category, the measurements of Boyne and Franken[33] and of Mamyrin and Frantsuzov[34] are also shown, though we did not use them in our adjustment. The steepest

50 MHz

portability

with Dual-Trace and Sweep Delay

Here's the new portable oscilloscope for DC-to-50 MHz applications. It operates almost anywhere —and under severe environmental conditions. it's small and light—with overall dimensions of 7¼" high x 12½" wide x 22½" deep (including extended carrying handle), and weighs less than 29 pounds.

Performance features include:

Bandwidth (with new P6010 Probe)
20 mV/div through 10 V/div > 50 MHz
10 mV/div > 45 MHz
5 mV/div > 40 MHz
1 mV/div > 25 MHz (Channels cascaded)

Sweep Rates—5 sec/div to 10 nsec/div (with 10X Mag.)

Calibrated Sweep Delay—50 sec to 1 μsec.

CRT—New 4" rectangular, operating at 10 kV.

X-Y Operation—DC to > 5 MHz, 5 mV/div through 10 V/div.

Triggering—To 50 MHz, from Channel 1 or combined signals (both sweeps).

Type 453 Oscilloscope U.S. Sales Price f.o.b. Beaverton, Oregon . . . **$1,950**

new 453 from tektronix

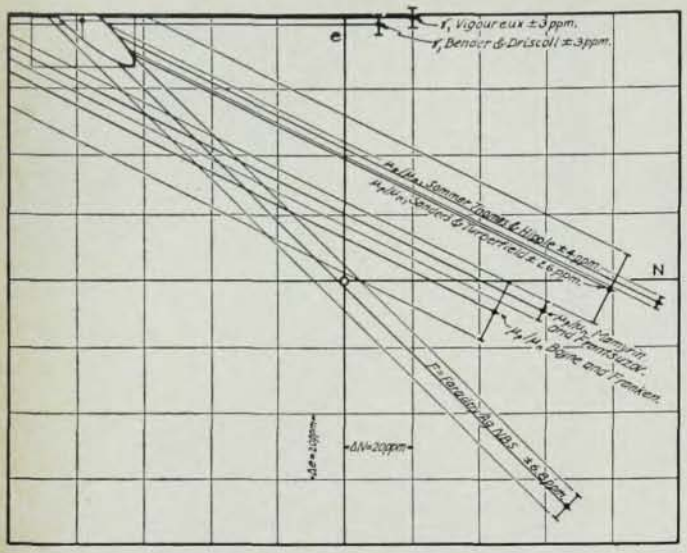


Fig. 11. Consistency graph for input data of 1963 adjustment if alpha is increased by 26 parts per million. This new value is implied by Refs. 27 and 28. This is thus a cross section of constants space in the plane on which alpha is 7.29739 × 10^{-3}. Central origin is the same as in Fig. 10, but the data imply an increase in e of 78 ppm and a decrease in N of the same relative magnitude. Consistency is the same as it was before.

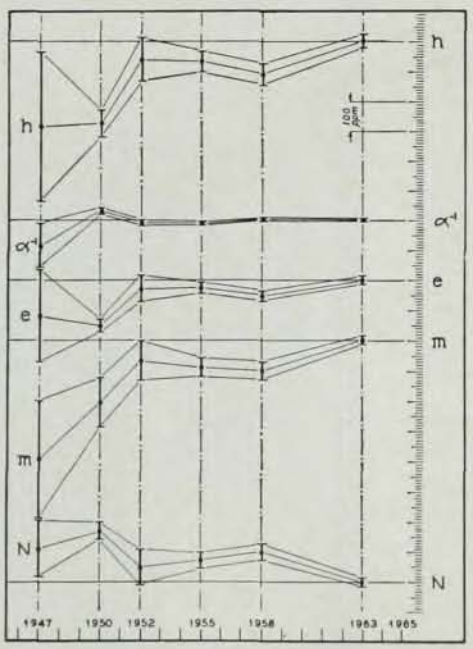


Fig. 12. History of fluctuations in five constants. Value and estimated standard deviation are shown from 1947 to 1963.

line is the most recent, very accurate, re-measurement of the Faraday constant at the National Bureau of Standards.[35] Figure 11 shows what happens to this diagram if we adopt the value of α calculated from the hyperfine splitting in hydrogen with the uncertain proton-structure correction terms. The diagram is the same as before, save that the ellipse of error and the entire complex of lines moves to higher values of e and lower values of N by 3 times 26 ppm.* The crucial nature of our choice of α thus becomes evident. R. T. Robiscoe, now at Yale, has already remeasured the Lamb shift, $2^2S_{1/2}-2^2P_{1/2}$, by observing level crossings in hydrogen, and has obtained a result differing significantly from Lamb's.[36] The Lamb shift is only a *part* (roughly one-tenth) of the total splitting, $2^2P_{3/2}-2^2P_{1/2}$, on which α depends, but if the remainder, $2^2P_{3/2}-2^2S_{1/2}$, does not change at all, and Robiscoe's change in the Lamb shift remains a solid fact, that alone will suffice to change α upward by roughly the 26 ppm needed for agreement with the hyperfine value. This unsolved problem of α, however, is one which I must bequeath to my younger colleagues.

The graph of Fig. 12 shows a history of the fluctuations in value and in assigned precision of five of the constants in our five adjustments from 1947 to 1963. On this graph the smallest vertical divisions correspond to changes of ten parts per million. In 1947 the errors were of order 200 ppm, now they are of order 10 or 20 ppm.

Here, then, are the crooked paths that, since 1947, Dr. Cohen and I, in partnership, have been treading as pilgrims in search of the constants. Each succeeding adjustment has settled some question, but has opened up others, equally fascinating but on a higher level of precision. We believe that this has served a useful purpose by calling attention from time to time to the trouble spots where further research was needed. The level of precision is still some distance from approaching comparability with the accuracy and reproducibility of our primary standards defining length and mass. The precision of definition of time with the hydrogen maser of Dr. Ramsey and his group at Harvard is now, of course, still many orders of magnitude superior. I think further progress with

* This statement is based on the form of the six final equations of observation of the 1963 adjustment (Refs. 24 and 25) which fall into the following four kinds:—

$$\alpha = C_1 \quad \text{(Refs. 26 or 27)} \quad (1)$$
$$Ne^2\alpha^{-3} = C_2 \quad \text{(Refs. 31 and 32)} \quad (2)$$
$$Ne = C_3 \quad \text{(Ref. 35)} \quad (3)$$
$$\alpha^3e^{-1} = C_4 \quad \text{(Refs. 29 and 30)} \quad (4)$$

wherein the C's are stochastic numerics depending only on the results of the physical measurements referred to above by bibliographical reference number and on "auxiliary constants". By the latter is meant constants such as c, the speed of light, or $R\infty$ the Rydberg for infinite mass, now so accurately known that their uncertaintities contribute negligibly to the error of the C in question.

the constants from here on is going to be slower and more painful, since on the present level of accuracy the experiments take much longer, cost more, and are, therefore, not duplicated by many people. I cannot emphasize too strongly, however, the importance of much more wide-spread duplication, using many different approaches by many different groups, because here we are dealing with the foundations of physics upon which all the rest of the superstructure is based. I think the spirit of our times presents a grave danger that we may attach so much undeserved importance to mere standardization and conformity of usage that we lose sight of the fact that scientific truth in this field is only to be reached by looking for discrepancies and disagreements. It is a free battlefield, no place for worshippers of the cult of Dale Carnegie, I assure you.

Near the close of John Bunyan's famous book, *Pilgrim's Progress*, Mr. Valiant-for-Truth, to whose exemplary virtues I make no claim but whose parting words I shall borrow, says: "Though with great difficulty I am got hither, yet now I do not repent me for all the trouble I have been at to arrive where I am. My sword I give to him that shall succeed me in my pilgrimage and my courage and skill to him that can get it. My marks and scars I carry with me. . . ."

References

1. J. A. Bearden, Phys. Rev. **37**, 1210 (1931); J. Appl. Phys. **12**, 395 (1941).
2. E. Bäcklin, Z. Physik **93**, 450 (1935).
3. M. Söderman, Nature **135**, 67 (1935); Dissertation, Uppsala (1934).
4. F. Tyrén, Z. Physik **109**, 722 (1938); Dissertation, Uppsala (1940).
5. J. W. M. DuMond and V. L. Bollman, Phys. Rev. **50**, 524 (1936).
6. J. W. M. DuMond, Phys. Rev. **56**, 153 (1939).
7. J. W. M. DuMond, Phys. Rev. **58**, 457 (1940).
8. J. W. M. DuMond and E. R. Cohen, Revs. Modern Phys. **20**, 82 (1948).
9. J. W. M. DuMond and E. R. Cohen, report to the National Research Council Committee on Constants and Conversion Factors of Physics: "A Least-Squares Adjustment of the Atomic Constants as of December 1950" (unpublished).
10. E. R. Cohen, J. W. M. DuMond, T. W. Layton and J. S. Rollett, Revs. Modern Phys. **27**, 363 (1955).
11. A. Petermann, Helv. Phys. Act. **30**, 407 (1957); C. M. Sommerfield, Phys. Rev. **107**, 328 (1957); E. R. Cohen and J. W. M. DuMond, Phys. Rev. Letters **1**, 291 (1958).
12. A discussion of error statistical-correlation is given in Chapter 7 of E. R. Cohen, K. M. Crowe, and J. W. M. DuMond, *Fundamental Constants of Physics* (Interscience Publishers, Inc. New York, 1957). For the generalized formula of error propagation see Section 7.4.
13. I. Henins and J. A. Bearden, Phys. Rev. **135**, A890 (1964).
14. J. A. Bearden, Phys. Rev. **38**, 2089 (1931); A. Ievins and M. E. Straumanis, Z. Physik **116**, 194 (1940).
15. J. A. Bearden, "X-Ray Wavelengths", Technical Report No. NYO-10586, AEC Division of Technical Information, Oak Ridge, Tenn., 1964 (unpublished).
16. J. W. Knowles, Can. J. Phys. **40**, 257 (1962); *Proceedings of the Second International Conference on Nuclidic Masses*, Vienna, July 1963 (Springer-Verlag, Vienna, 1964).
17. S. Hagström, Dissertation, Uppsala (1964); S. Hagström, O. Hornfeldt, C. Nordling and K. Siegbahn, Arkiv Fysik **23**, 145 (1962).
18. J. J. Spijkerman and J. A. Bearden, Phys. Rev. **134**, A871 (1964).
19. J. A. Bearden, Phys. Rev. **137**, B181 (1965).
20. A. Smakula and V. Sils, Phys. Rev. **99**, 1744 (1955); A. Smakula, J. Kalnajs, and V. Sils, Phys. Rev. **99**, 1747 (1955).
21. B. Edlén and L. K. Svensson, Arkiv Fysik **28**, 427 (1965).
22. J. W. M. DuMond and E. R. Cohen, Phys. Rev. **103**, 1583 (1956); J. W. M. DuMond, Proc. Nat. Acad. Sci. **45**, 1052 (1959).
23. These tests are explained in both references 24 and 25 under the heading "Measures of Incompatibility of an Input Datum".
24. E. R. Cohen and J. W. M. DuMond, report to National Research Council Committee on the Fundamental Constants of Physics and Chemistry: "An Analysis of Variance of the Data Available in 1962" Feb. 28, 1963 (unpublished). Copies available from A. G. McNish, Committee Chairman, National Bureau of Standards, or from authors.
25. E. R. Cohen and J. W. M. DuMond, "Present Status of our Knowledge of the Numerical Values of the Fundamental Physical Constants," *Proceedings of the Second International Conference on Nuclidic Masses*, Vienna, July 1963 (Springer-Verlag, Vienna 1964). Copies obtainable from authors.
26. S. Triebwasser, E. S. Dayoff, and W. E. Lamb, Jr., Phys. Rev. **89**, 98 (1953).
27. S. B. Crampton, D. Kleppner and N. Ramsey, Phys. Rev. Letters **11**, 338 (1963).
28. C. K. Iddings, Phys. Rev. **138**, B446 (1965).
29. P. L. Bender and R. L. Driscoll, Trans. I.R.E. **I-7**, 176 (1958).
30. P. Vigoureux, Proc. Roy. Soc. (London) **A270**, 72 (1962).
31. H. Sommer, H. A. Thomas and J. A. Hipple, Phys. Rev. **82**, 697 (1951).
32. J. H. Sanders and K. C. Turberfield, Proc. Roy. Soc. (London) **A272**, 79 (1962).
33. H. S. Boyne and P. A. Franken, Phys. Rev. **123**, 242 (1961).
34. B. A. Mamyrin and A. A. Frantsuzov, Dokl. Akad. Nauk SSSR **159**, 777 (1964).
35. D. N. Craig, J. I. Hoffman, C. A. Law and W. J. Hamer, Nat. Bur. Std. J. Res. **64A**, 381 (1960).
36. R. T. Robiscoe, PhD thesis in Physics, University of Chicago, 1964; present address of author, Physics Department, Yale University.