Mathematical Analysis of the Universal Physical Constants.

E. R. COHEN

Atomics International
A Division of North American Aviation, Inc., Canoga Park, California
and California Institute of Technology - Pasadena, California

In the previous paper Professor DuMond has presented a wealth of experimental data on the universal physical constants (primarily those known as the atomic constants). This material is made up of a score or more individual experimental results. Not all of these are of equal experimental accuracy. Some of them may be combined among themselves to yield a simple datum which represents an average value of several independent evaluations of the same physical quantity. In whatever way this may be done we still find ourselves with more data than are required to obtain an evaluation of the fundamental physical constants. This may be expressed by a classical example (the example is classical in that it represents one of the first instances in which the existence of the overdetermined nature of the experimental data became forcefully evident.)

The value of the faraday constant as measured by the silver voltameter is [1] $F = Ne = 9651.29 \pm 0.19$ emu/mole; the atomic weight of the electron as inferred from the isotopic shift in the Balmer series for hydrogen and deuterium is [2] $Nm = (54.895 \pm 0.008) \cdot 10^{-5}$ amu. From this we can calculate a value for the specific electronic charge

$$e/m = (1.75814 \pm 0.00026) \cdot 10^7 \text{ emu/g} .$$

In comparison with this the value of e/m determined from measurements of the acceleration and deflection of electrons in electromagnetic fields yields a value [3]

$$e/m = (1.75959 \pm 0.00038) \cdot 10^7 \text{ emu/g} .$$

We could of course have equally well used the measured values of the atomic weight of the electron and the specific electronic charge to calculate an « indirect » value of the Faraday constant, or conversely use the specific electronic charge and the Faraday to find an indirect value of the electron's. atomic weight. These calculations are summarized in Table I. We are then.

TABLE I. – *Comparison of direct and indirect values from a simple overdetermined system.*

	$F = Ne$ (emu/mole)	Nm (amu)	e/m (emu/g)
Direct Values	$9\,651.29 \pm 0.19$	$(54.895 \pm 0.008) \cdot 10^{-5}$	$(1.759\,59 \pm 0.000\,36) \cdot 10^{7}$
	$(Nm)(e/m)$	$F/(e/m)$	F/Nm
Indirect Values	$9\,659.32 \pm 1.83$	54.849 ± 0.006	$1.758\,14 \pm 0.000\,26$

presented with the embarrassing problem of choosing which consistent set of three numbers ought to be used; each choice is certainly poor in itself since each is contradicted by the other two. It should be emphasized that the numbers presented in this table are illustrative and although based on direct measurements, do not represent in any way a selection of « best » data or are they to be considered as recommended numbers. They are presented in order to show the nature of the difficulty (which is actually compounded many-fold) when we attempt to analyze the experimental data on the atomic constants, considering these data as a single unified complex of experimental results. The experiments described in the preceding paper measure various combinations of the atomic constants. In some experiments the same combination or the same constant is measured in different ways; in others a measurement is made of a quantity whose value is also deducible from combining the results of two or more different and independent experiments. Thus we are confronted with an inter-related complex which is similar (in an abstract form) to the situation which is found most often in geodetic triangulation. Let us therefore look more closely at the structure of our present problem .

Each of the quantities whose measurement is described in the preceding paper can be expressed in terms of h, e, m, N, Λ, c, and a few other auxiliary quantities whose numerical values are accurately known. The Rydberg constant, however, which is a function of h, e, m, and c, is much more accurately measured than are any of its component factors. It is therefore convenient to consider the Rydberg constant as an exact numerical quantity and to use it to express the electron/mass as a known function of the other variables. In addition, to the accuracy required in the discussion to follow, it is convenient to treat the velocity of light, c, as an exactly known quantity. Furthermore,

because of the fact that the fine structure constant, α, is more accurately determined by direct observation of fine structure in hydrogen than it can be computed by combining, say measurements of e, h/e and c, it is also convenient to use α as a variable in place of h. Thus we shall choose to express all of the experimental results in terms of the quantities α, e, N, and Λ which we may call the primary unknowns of our analysis. The result of any experiment can then be described as measuring (except for quantities which can be considered as accurately calculable correction factors) some product of powers of the primary variables of the form

$$(1) \qquad\qquad f(\alpha, e, N, \Lambda) = \alpha^a e^b N^c \Lambda^d = A(1 \pm \sigma) \,.$$

The second expression is not actually an equation; it is a short-hand method for indicating that we do not know the true value of the right hand side of the equality but that the experiment can be interpreted as indicating a value which is defined only to the extent of a probability distribution with mean value A and with relative standard deviation σ. This interpretation is not the only one however and we shall discuss a different interpretation later.

For convenience in analyzing the data it is useful to linearize our equations. We adopt origin values α_0, e_0, N_0, Λ_0, which have been chosen sufficiently close to our expected solution that any set of values α, e, N, Λ (any point in our « constants-space ») in which we are likely to be interested will differ from the origin values by small amounts. « Small » is determined here by the magnitudes of the second derivatives of the function f; we assume that we can expand $f(e, \alpha, N, \Lambda)$ as a multivariate Taylor series about the origin point, and that in such an expansion only the linear terms need be retained. It is in practice most convenient to expand the logarithm of f in a Taylor series; we then deal with relative deviations and if the function is a simple product of powers of the variables as is the case here, we obtain, in place by Equation (1), the linearized equation

$$(2) \qquad\qquad a x_\alpha + b x_e + c x_N + d x_\Lambda = h \pm \sigma,$$

where $x_\alpha = (\alpha - \alpha_0)/\alpha_0$; $x_e = (e - e_0)/e_0$, etc., and $h = (A - A_0)/A_0$ in which A_0 is the value of the function f evaluated with the origin values α_0, e_0, N_0, Λ_0.

1. – A geometrical interpretation.

There are various ways in which we can interpret the experimental data geometrically. Professor BIRGE in his paper has discussed various types of diagrams [4-7] which he has used to aid him in evaluating « best » or « most

consistent » values of the atomic constants whose numerical magnitudes he seeks. These geometrical descriptions are all representations (in terms of various types of projections) of the following geometrical structure.

We may represent each experimental determination as defining a single functional relationship among the various physical constants. We may consider that an independent set of constants forms an orthogonal co-ordinate system in a « constants-space ». Each functional relationship (each experiment) then is represented by a surface in this space. An arbitrary point in the space (a given set of values of the constants, say α, e, N, Λ) is in agreement with the experimental result if it lies in the surface. A point (that is, a set of values for the constants) is therefore consistent with several experimental results if it lies at the common intersection of the surfaces. In general if we have more experiments (surfaces or equations) than we have constants to be determined (dimensionality of the space) it is not a priori evident that there will be a point through which *every* surface passes. If the experiments are reasonably consistent there will be a point, however, which is close to all of the surfaces. The best choice for the set of physical constants is that point which is « closest » to all of the surfaces. The definition of « distance » in this space (that is, the definition of a metric) we shall discuss below; for the moment let us *assume* that we can define distance. We should then want to consider a metric function defined with regard to the point in question and the existing surfaces. The problem we then set for ourselves is to find the point which minimizes this function.

This leads us to a second geometrical representation of the problem. At each point in our « constants-space » we have defined the value of a certain function. We now consider a space of one higher dimensionality and consider a hyper-surface whose distance from an orthogonal subspace (the original constants-space) is defined at each point by the value of a certain function (the exact form of this function we have not yet specified). We then seek the point on this surface which is a minimum. It is possible to make several general remarks about the form of the function, which we shall call Q. It must primarily be defined in such a way that its value is independent of the actual variables used to describe it. This means that Q is a scalar with respect to a transformation or rotation in the « constants-space ». In addition it should also be invariant with respect to non-essential modifications in the form in which the experimental results are presented. Again a simple example may be in order.

Let us suppose that one of our experimental equations can be written in the form

(3) $$ax_1 + bx_2 + cx_3 = h \pm \sigma,$$

which states that the experimental result h was obtained with a standard

error $\pm \sigma$. If we transform variables from x_1, x_2, x_3 to a new set of variables y_1, y_2, y_3, we obtain in place of this equation the expression

$$(3a) \qquad\qquad \alpha y_1 + \beta y_2 + \gamma y_3 = h \pm \sigma .$$

The function Q must be defined in terms of the coefficients a, b, c, (and the corresponding coefficients of all of the other experimental data) in such a way that $Q(\alpha, \beta, \gamma; y_1, y_2, y_3)$ should be equal to $Q(a, b, c; x_1, x_2, x_3)$. Furthermore Q must be invariant with respect to tranformation of the equations; the most easily recognized type of function transformation would be the replacement of equation (3) by the expression]

$$(3b) \qquad\qquad nax_1 + nbx_2 + ncx_3 = nh \pm n\sigma .$$

When this is done it is perhaps suggestive that all of the experimental equations ought to be written in a normalized form

$$(4) \qquad\qquad \frac{a}{\sigma} x_1 + \frac{b}{\sigma} x_2 + \frac{c}{\sigma} x_3 = \frac{d}{\sigma} \pm 1 .$$

When the equations are written in this form all of the equations have equal accuracy and we might then infer that they ought to be given equal weight in the analysis. At this point we use the word « weight » in a qualitative sense; later we shall give a quantitative definition which will formalize the intuitive notion that data with small standard errors should be given more weight or importance and data with larger errors should be given less weight in any analysis.

2. – The criterion of least squares.

Until now we have purposely made only very general statements about the problem of finding the best or most consistent set of unknowns from a set of overdetermined equations. It may perhaps be apparent that we would like to introduce the criterion of least squares in a logical and completely deductive manner. It is almost certain that this cannot be done; we can however carry our deductive development to a certain point, then inductively introduce the least squares criterion and finally justify the choice (at least partially) by considering the conditions which it must satisfy. We could specify that the function Q which has been previously mentioned shall be defined as the sum of the squares of the deviation of the point from each of

the experimentally defined surfaces, each deviation normalized by dividing it by the standard deviation of the experimental measurement. The condition that Q be a minimum is therefore the usual condition of « least squares » which was first formulated by GAUSS about 1820 [8].

There have been many who have objected to the use of the method of least squares on the grounds that it assumes the Gaussian curve for the probability distribution of the error in any particular measurement. The objection is indeed valid that the assumption of the Gaussian distribution is often unwarranted in many experimental configurations. The real problem at issue, however, is one of determining a « best » set of values that can be computed from an over-determined system of equations, and this is essentially the problem of determining an analytic basis on which one can define the adjective « best. » The condition of least squares serves as one such analytic criterion. This says nothing in itself of what could be called the physical interpretation of the criterion. It is recognized, in general, that the method of least squares corresponds to the « Axiom of Maximum Likelihood », if the distribution functions of all the errors are Gaussian [9]. GAUSS himself was able to justify the method on a much wider base, and in 1821 he published a theory which replaces this axiom with an « Axiom of Minimum Error » or « Axiom of Maximum Weight. » [10]. The definition of « best » is not to be made on the basis of that solution which is most likely to be correct, but on that combination of data which yields the most accurate result and to which can, therefore, be attached the greatest statistical weight. (The statistical weight of a statistical variable is defined in the present sense as the reciprocal of the variance, i.e., the reciprocal of the mean square error, of the quantity.) Consider then an overdetermined set of n equations expressing relationships between q variables. We can find an infinite number of solutions for the q variables depending on how we choose to combine the equations. We can think of this process as one in which some set of $q-1$ of the equations are used to express $q-1$ of the variables in terms of one particular variable, say x_1. These values are then to be substituted into the remaining $n-q+1$ equations to give a set of $n-q+1$ values for the variable x_1. Any of these numerical values is a possible choice for the variable x_1 and, in general we would want to take some weighted average of these numerical values. The numerical value of such an average depends both on the weights attached to the elements that go to make up the average and on the numerical values of these elements. Ultimately, therefore, the value ascribed to x_1 depends on $n-q$ independent and arbitrary parameters which specify the mathematical form of the average and n quantities which are either the observational numerics or quantities directly deduced from them (such as relative deviations from a set of origin values) which determine the numerical value of the average. There are only $n-q$ arbitrary parameters rather than $n-q+1$ because the

choice of the $n - q + 1$ weights is restricted by the condition that their sum is unity.

The numerical value of x_1 is thus expressed as a linear combination of n numerical quantities, each of which has associated with it a mean square error. If the numerical quantities are observationally independent, we can assert that the mean square error of x_1 is the sum of certain coefficients depending on the $n - q$ free parameters, times the mean square errors of the n observational numerics. The Axiom or Condition of Minimum Error states that the « best » choice for x_1 is that one whose error is a minimum with respect to the possible variation of the free parameters. This condition is equivalent to the condition of least squares, although the results cannot, in general, be identified as corresponding to that set which has maximum likelihood except in the case when the distribution function for the errors is specified to be Gaussian. This is however, an advantage for one can easily construct distributions (for example, rectangular distributions) for which the condition of maximum likelihood has no unique solution. Furthermore, the development of the condition of minimum error is quite general in regard to the forms of the error distribution functions; all that is specified is the mean square error, so that the range of applicability of the theory of least squares is extended from Gaussian distributions to the much larger class of distributions with finite second moments.

As an elementary example of the method and in order ot clarify the concepts involved, let us consider a problem which is perhaps the simplest possible example. We have two measurements of the quantity x; these two measurements are a_1 and a_2, in general, $a_1 \neq a_2$. We also assume that each measurement represents a single selection from a universe of values. We let the probability distribution of the first measurement be $P_1(\xi)$ such that the probability is $P_1(\xi) \, d\xi$, that the result of a measurement of the quantity x by the specified procedure shall lie between the values ξ and $\xi + d\xi$. Similarly, the second measurement of x is to be characterized by the probability distribution $P_2(\eta)$. We need not make any detailed specification of the form of the distribution functions P_1 and P_2; it is not even necessary that the two functions have similar form. We impose upon them only the restriction that the distributions have finite second moments.

If we have two measurements of x then we can take some average value to use as the « best » choice. There are, however, several different choices for this average and we can write

(5) $$x_0(\alpha) = \alpha a_1 + (1 - \alpha) a_2 \,,$$

where α is any real number, although intuitively we would prefer that $0 < \alpha < 1$ because this is the condition that x_0 lies between the values a_1 and a_2. Defined in this way, the mean of the universe of values of x_0 is

$$(6) \qquad \bar{x}_0 = \iint [\alpha\xi + (1 - \alpha)\eta] P_1(\xi) P_2(\eta) \, \mathrm{d}\xi \, \mathrm{d}\eta \, .$$

Each probability distribution is assumed to be normalized and if there are no systematic errors in either measurement, the expectation value of each measurement is x.

Therefore we find

$$(7) \qquad \bar{x}_0 = \alpha x + (1 - \alpha)x = x \, ,$$

so that the expectation value of the average is indeed the quantity we are trying to measure, independent of the parameter α which determines the particular average. But we now ask the question: « How accurate is this average; what is its standard deviation? » If we let ε^2 be the mean square deviation of the universe from which x_0 is extracted, we have, by the usual definition,

$$(8) \qquad \varepsilon^2 = \iint [\alpha\xi + (1 - \alpha)\eta - x]^2 P_1(\xi) P_2(\eta) \, \mathrm{d}\xi \, \mathrm{d}\eta =$$

$$= \alpha^2 \int (\xi - x)^2 P_1(\xi) \, \mathrm{d}\xi + (1 - \alpha)^2 \int (\eta - x)^2 P_2(\eta) \, \mathrm{d}\eta \, .$$

The two integrals in the last form of Equation (8) are the variances (the mean square errors), respectively, of the first and of the second measurements. These quantities are defined to be σ_1^2 and σ_2^2. The expression for the error in the average in terms of the errors of the numbers entering into the average is, therefore,

$$(9) \qquad \varepsilon_0^2 = \alpha^2 \sigma_1^2 + (1 - \alpha)^2 \sigma_2^2 \, .$$

We find that, although the expectation value of the average is independent of α, the error in the average *is* a function of α. For $\alpha = 0$ we have $\varepsilon^2 = \sigma_2^2$ and for $\alpha = 1$ we have $\varepsilon^2 = \sigma_1^2$; we may reasonably ask whether a proper choice of α might not result in a value of ε^2 which is smaller than either of these and indeed what value of α leads to a minimum value of ε^2. It is easily shown that the minimum value of ε^2 is

$$(10) \qquad \varepsilon_0^2 = \sigma_1^2 \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$$

which is achieved when $\alpha = \sigma_2^2 / (\sigma_1^2 + \sigma_2^2)$ and the corresponding value of x_0 is

$$(11) \qquad x_0 = \varepsilon_0^2 \left[\frac{a_1}{\sigma_1^2} + \frac{a_2}{\sigma_2^2} \right] .$$

The « weights » to be attached to measured quantities in order to compute the « best » average are therefore proportional to the reciprocal of the mean square error of each measurement. If we define the statistical weights

$$w_1 = 1/\sigma_1^2 \, , \qquad w_2 = 1/\sigma_2^2 \, ,$$

we have two important formulas

(12a) $$x_0 = (w_1 a_1 + w_2 a_2)/(w_1 + w_2) \, ,$$

(12b) $$w_0 = 1/\varepsilon_0^2 = w_1 + w_2 \, .$$

The first of these essentially justifies the nomenclature of « statistical weight » for the reciprocal variance, since it is this quantity that determines the importance of the measurement in the computation of the average. The second formula shows that when this « best » average is obtained the weight of the result (computed as the reciprocal of the variance of the average) is just the sum of the weights of the components and, furthermore, that this is the maximum weight ascribable to any average. An other linear combination of the two observations would have a weight which is less than the sum of the individual weights and is, therefore, an inefficient average that wastes weight. The statistical weight of the average is a maximum (equal to the sum of the weights of the individual components), when the weighting employed in computing the average is determined by the statistical weights of the components.

3. – A generalized theory of least squares.

The example which we have just considered can be immediately generalized to consider the combination of several measurements instead of only two. Furthermore it is also rather straightforward (although lengthy) to extend the concepts introduced above into the problem of finding the best set of values of several simultaneous variables related by linear equations. This is of course the sort of system in which we are interested once we have linearized the equations of observation as previously indicated.

We write the linearized system in the form

(13)
$$\begin{cases} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1q}x_q = c_1 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2q}x_q = c_2 \\ \ldots\ldots \quad \ldots\ldots \quad \ldots\ldots \qquad \ldots\ldots \quad \ldots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nq}x_q = c_n \end{cases}$$

We shall use Roman letters to indicate the indices which range from 1 to q and which refer to the unknowns, and Greek letters for those which range from 1 to n and refer to the observational equations.

Since the c_μ are assumed to be independent, the errors η_μ in c_μ are uncorrelated and

(14)
$$\begin{cases} \langle \eta_\mu \eta_\nu \rangle = \sigma_\mu^2 & \text{if } \nu = \mu, \\ \qquad\quad = 0 & \text{if } \nu \neq \mu. \end{cases}$$

We wish to find x_i as a linear combination of the c with appropriate coefficients, λ_μ^i chosen so as to determine the x_i with maximum accuracy or minimum error. Thus we write

(15)
$$x_i = \lambda_1^i c_1 + \lambda_2^i c_2 + \lambda_3^i c_3 + \dots + \lambda_n^i c_n$$
$$= \sum_\mu \lambda_\mu^i c_\mu = \sum_{\mu j} \lambda_\mu^i a_{\mu j} x_j \, .$$

This equation is to be understood as follows. Each of the n individual equations of observation (identified by the index μ is to be multiplied by a number, λ_μ^i, there being a set of n such numbers for each variable x_i. These numbers are to be chosen such that, when the n equations are then summed, the coefficient of each x, other than the specified x_i, is zero, while the coefficient of x_i is unity. This implies that there are q conditions on each of the q sets of n multipliers. This is not sufficient to define the multipliers since $n > q$. The additional conditions are obtained by requiring that we choose that linear combination for each variable x_i which will have the smallest variance and hence the largest possible weight. It can be shown [11] that such a condition leads to exactly the same values of x_i as would be obtained from the condition that we minimize the quadratic form

(16)
$$Q = \sum_\mu \left(\frac{a_{\mu 1} x_1 + a_{\mu 2} x_2 + \dots + a_{\mu q} x_q - c_\mu}{\sigma_\mu} \right)^2 .$$

Hence the condition of Minimum Error is just the least squares condition and is therefore equivalent to the condition of Maximum Likelihood of the more restricted Gaussian case. It is, however, much more general than the latter. No restriction is put on the probability distributions of the input errors, η_μ, in the c_μ other than that these distributions must have finite second moments, i.e., the σ_μ must exist. Subject to this condition, they may have more than one « mode » or indeed any arbitrary shape; nor is it necessary that the distribution functions all have the same form.

This generality of the significance of the least squares adjustment when stated in terms of the second moments of the error distribution is of great importance. It emphasizes also the desirability of adopting the root-mean-square deviation as a measure of error in preference to such error measures as the « probable error », or the « mean absolute error. » For a gaussian distribution, the three measures, root-mean-square error, mean absolute error and probable error (that error which divides the distribution curve into equal areas, so that the probabilities of errors of absolute magnitude greater than or less than the probable error are equal) stand in the ratios $1:0.798\ 0:0.674\ 5$. It is however a mistake to think of the different error measures as simply expressing the same error spread on different scales. When we do not limit ourselves to gaussian distribution the root-mean-square error or standard error enjoys a position of far greater statistical significance and generality than do the others. This is because the standard deviation has a simple reproduction property (for any form of probability distribution) which is not shared by any other parameter with the same generality.

It may even be worthwhile to avoid the discussion of probability distributions and variances entirely and speak only of statistical weights and the weight to be assigned to each experimental number or to any number that results from arithmetic combinations of experimental data.

It can be shown, independently of the development which led to Equation (16) that if the weight to be assigned to a random variable is to be in accord with certain elementary axioms regarding the definition of weight and at the same time be a function only of the standard error of the variable, then that function is uniquely determined to be the reciprocal of the variance. The weight of a quantity should therefore be defined in such a way that it is consistent with the identification as the reciprocal of the variance in those cases where a variance can be experimentally determined. But we should be prepared as well to admit to consideration subjective evaluations of the weight of an experiment based on an impartial and honest appraisal of the inherent accuracy of one experiment in comparison to another even though no analytic evaluation of the variance exists. It may well be that the introduction of « weights » in place of « variances » is purely a semantic subterfuge; even so, it may prove a useful concept.

If, in equation (16), we introduce the notation

$$(17) \qquad r_\mu = a_{\mu 1} x_1 + a_{\mu 2} x_2 + \dots + a_{\mu q} x_q - c_\mu ,$$

we can write Q in a matrix or tensor form [11, 12]

$$(18) \qquad Q = r_\mu \pi_{\mu\nu} r_\nu = R^+ \Pi R$$

in which $\pi_{\mu\nu}$ are the elements of a diagonal matrix

(19)
$$\begin{cases} \pi_{\mu\nu} = 1/\sigma_\mu^2 & \mu = \nu, \\ \quad\;\; = 0 & \mu \neq \nu, \end{cases}$$

R is the vector with components r_μ and R^+ is the transposed vector. This form may appear cumbersome and unvieldy in actual use, and it is if the simpler formulation is available. However, it represents the point of departure for generalizing the formulation and establishing its complete invariance. If we perform a linear transformation on R and obtain $R' = TR$ we must write

(20)
$$Q = R'^+ T^{-1+} \Pi T^{-1} R',$$

since Q must be a scalar invariant. Hence the transformed weight matrix Π' is no longer diagonal. It now has the form

(21)
$$\Pi' = T^{-1+} \Pi T^{-1} = [TST^+]^{-1},$$

where S is the inverse of the diagonal matrix Π; the diagonal elements of S are the variances σ_μ^2. We see immediately however that the transformed matrix $S' = TST^+$ is exactly the matrix which now expresses the variances and covariances of the transformed vector components r_ν'. The elements of S' are the mean error products, $\langle \eta_\mu' \eta_\nu' \rangle$. In the transformed system the errors η_μ' are not in general independent and hence S' is not diagonal as is the error matrix in equation (14). Thus, the form given in equation (18) represents the generalized statement of the least squares condition; the weight matrix Π is the inverse of the error matrix, or covariance matrix, of the observational date. If the observational equations are not independent, the covariance matrix S and the weight matrix Π are not diagonal. The off-diagonal elements of S are directly related to the degree to which the corresponding observational data are interdependent.

4. – Standard errors and correlation coefficients.

The statistical errors to be assigned to the output values of any least squares adjustment must, in general, be described not only by stating the standard deviation for each numerical result but also by specifying the correlations which exist between each pair of results. The numerical output values of α, e, N, etc., are of little use unless functions of these quantities can be com-

bined to compute other derived values. The standard deviations of such derived values must be computed by formulae which involve not only the standard
deviations of the values entering into the function but also the covariances v_{ij}
connecting all possible pairs of those values. This is because the output values
of a least squares adjustment are not in general statistically independent quantities but are « statistically correlated ».

Each quantity subject to random or accidental error (frequently known
as a «random variable ») may conveniently be thought of as a sample taken
at random from a « universe » of values which group themselves around a
mean value according to some frequency law. For each such random variable x_i, one is to think then of the implied universe of values from which the
variable is selected. This universe may be described by giving some of its
parameters. Thus, if the universe is known to be Gaussian, for example, then
prescription of its first and second moments, i.e., its mean value, μ_i, and its
variance σ_i^2 are sufficient.

Two such random quantities are observationally independent if the random
selection of a sample value from one universe in no wise affects or biases the
free selection of a sample from the other universe. If, for example, two
variables are connected by a strict functional relationship so that the value
of either one is uniquely determined by the other, the variables are completely
correlated and the correlation coefficient connecting them has the absolute
value unity. Random samples can no longer be selected freely and independently from the two universes because of the functional condition which
ties the selections rigidly together. Having selected a sample value from one
universe at random the second selection is now completely specified. On the
other hand, if one of the two random variables is a function of the other and
also of still other completely independent random variables, then these two
will be partially but not completely correlated, and the correlation coefficient
connecting them will have a value somewhere between -1 and $+1$. In
general if we have a set of statistically independent random variables, y_μ,
then a second set of variables, x_i, obtained by linear transformation on the y_μ

$$(22) \qquad\qquad x_i = \sum_\mu \lambda_{i\mu} y_\mu$$

will not be statistically independent since a given one of the variables y_μ and
hence a given source of error is present in more than one of the variables x_i.
This of course is just the situation which exists in the problem of least squares
adjustment. To each output value, x_i, there corresponds a standard deviation, σ_i, (« variance » $v_{ii} = \sigma_i^2$) and to each pair (x_i, x_j) there corresponds a
« covariance », $v_{ij} = r_{ij}\sigma_i\sigma_j$. The entire set of variances and covariances form
a symmetric matrix which we may call the « error matrix. » The elements

of this matrix are required in order to compute the error measures of other quantities depending on the x_i.

5. – The standard errors of the residues of a least squares adjustment [13].

In any least squares adjustment of data it is obviously important to be able to assign a standard deviation to the difference between the adjusted output datum and the input datum from which it was obtained. Closely related to this question is the question of the value which would be obtained in a least squares analysis if a specific input datum had been omitted. The input and output data are certainly correlated and it would be incorrect to calculate the standard deviation of the difference without considering this correlation.

The measured input data consist of numbers c_μ as in equation (13); after the least squares solution has been performed and we have obtained the adjusted values for the x's we can insert these values into equation (13) to obtain the adjusted value of c_μ which we shall designate by c_μ^*. The number c_μ^* is the best estimate which we can make of the correct value of c_μ based upon all the data available to us. The difference between c_μ and c_μ^* is thus a measure of the extent to which the observed value c_μ is consistent with all of the other data. In order to evaluate this consistency we must have an estimate of the magnitude of the difference which might be expected on the basis of statistical fluctuations. It can be shown fairly easily that the variance of the difference, $c_\mu - c_\mu^*$ is given by $\sigma_\mu^2 - \sigma_\mu^{*2}$ where σ_μ^2 is the variance of the input datum and σ_μ^{*2} is the variance of the adjusted value. This is a surprisingly simple result and it justifies the description of the adjusted value as being compounded of two terms; one is the direct input value c_μ while the other is an effective or indirect value which is determined by the combined action of all of the other data. This indirect value is the value of c_μ which would be deduced from a least squares analysis from which the directly observed datum had been omitted.

The indirect value and its variance are given by

$$(23) \qquad c_\mu^i = \frac{\sigma_\mu^2 c_\mu^* - \sigma_\mu^{*2} c_\mu}{\sigma_\mu^2 - \sigma_\mu^{*2}} = c_\mu^* + \frac{\sigma_\mu^{*2}}{\sigma_\mu^2 - \sigma_\mu^{*2}} (c_\mu^* - c_\mu) \,,$$

$$(24) \qquad \sigma_\mu^{i2} = \frac{\sigma_\mu^2 \sigma_\mu^{*2}}{\sigma_\mu^2 - \sigma_\mu^{*2}} \,.$$

This expression for the variance implies that the statistical weights, which are proportional to the reciprocal of the variances are related by the equation

$$(25) \qquad p_\mu^* = p_\mu^i + p_\mu \,, \qquad p = C/\sigma^2 \,.$$

If we use weights rather than variances, Equation (23) takes on the simple form

$$(26) \qquad c_\mu^* = \frac{p_\mu c_\mu + p_\mu^i c_\mu^i}{p_\mu + p_\mu^i} \,,$$

which is merely the statement that the least squares adjusted value is the weighted mean of the direct input value and the indirect value.

6. – Analysis of data.

We now at last come to the problem of making a specific analysis of a given set of experimental data in order to determine best values. We have formulated a procedure for doing this; we must now determine the data to which this procedure is to be applied. One must be especially careful not to apply the method of least squares blindly; it is not a substitute for careful selection of data. No provision exists in the method for identifying and isolating systematic error; the comparison of χ^2 with the theoretical probability table is useful in this regard but it is not definitive and can only indicate a probable existence of systematic error.

In a least squares fitting each datum is to be assigned a weight which is inversely proportional to its variance. Such a weighting is not arbitrary; it can be deduced directly from the form of the quadratic expression

$$Q = \sum (r_\mu/\sigma_\mu)^2 \,,$$

which we attempt to minimize. An observation with a large variance therefore carries little weight in determining the value of Q, and hence may be omitted without greatly affecting the result. There is, however, a more important reason for omitting data of low weight. When an experimenter designs his experiment he must carefully consider the possible presence of systematic error as well as the presence of random error. The random error of the final result can be reduced by duplication and repetition since these errors are different in each repetition; the systematic errors, on the other hand do not cancel out but remain. Now it is proper in an experiment to reduce any possible source of systematic error to a point where it may be of the order of, say, one tenth of the random error of a single observation. In this way the systematic error will be of the same order as, or smaller than, the random error of the final quoted result. However, it is neither feasible nor practical to do much better than this in the suppression of systematic error.

Thus an experiment with a quoted error which is large compared to another similar but more precise measurement may well be affected with a systematic

error which is large compared to the accuracy of the second experiment. This systematic error would then, a fortiori, be large compared to the accuracy which might be claimed for weighted mean of the two results. It would therefore be inappropriate to include the less precise observation in a weighted mean.

It is also necessary to reject data which suffer from serious systematic error even if (and perhaps, especially if) the data are precisely measured. Grounds for the suspicion of such systematic error my arise either from experimental sources (such as a reevaluation of the conditions under which the experiment was performed under such circumstances that corrections for such systematic errors can not be made to the existing data) or from theoretical sources (in which, as in the case of the hyperfine structure splitting in hydrogen the precision of the experiment is higher than the esixting state of the theory of the experiment can handle).

By the first criterion for rejection mentioned above almost all of the historically important early experiments, including many which were considered important as late as 1947 are excluded. Within their estimated precision ranges such measurements are not inconsistent with the later more precise results but they are relatively so much less accurate as to have quite negligible influence in the present least squares adjustment. As a result of this criterion very few measurements published prior to 1950 remain.

Professor DuMond has surveyed, in the previous paper, the important experiments which have a bearing on the values of the atomic constants. We shall now collect here those results which are to be used in an analysis of these constants.

(i) The *conversion factor* Λ, from the Siegbahn nominal scale of X-ray wavelengths (in X-units) to milliangstroms. We shall use the value recommended by Sir LAWRENCE BRAGG since it presumably supersedes the earlier value quoted by T. T. BIRGE (*):

(27) $$\Lambda = 1.002\,020 \pm 0.000\,030 \,.$$

(ii) The *Siegbahn-Avogadro number* N_s'. We use Birge's value converted to the physical scale

$$N_s' = N\Lambda^3 = (6.061\,79 \pm 0.000\,23)\cdot10^{23} \text{ mole}^{-1} \,.$$

(iii) The *fine structure separation in deuterium*, ΔE_D: The frequency

---

(*) However, very recent information concerning TYRÉN's measurements on X-ray wavelengths indicates that this work suffers a systematic error of the order of 30 to 50 ppm because of the ommission of the Lamb shift in the calibration of the photographic plates. (See the previous paper by J. W. M. DuMond).

separation between the levels $2^2P_{\frac{3}{2}}$ and $2^2P_{\frac{1}{2}}$ measured by DAYHOFF, TRIEB-WASSER and LAMB

(29) $\varDelta E_\mathrm{D} = \dfrac{1}{16}\,\alpha^2\,R_\mathrm{D}c\left(1 + \dfrac{5}{8}\,\alpha^2\right)\left[1 + \dfrac{\alpha}{\pi} - 5.946\,\dfrac{\alpha^2}{\pi^2}\right] = (110\,971.59 \pm 0.10)\ \mathrm{MHz}\,.$

(iv) The *gyromagnetic ratio of the proton*, γ_p, obtained at the National Bureau of Standards by THOMAS, DRISCOLL and HIPPLE

(30) $\gamma_\mathrm{p} = \mu'Ne/M_\mathrm{p}c = (26\,725.3 + 0.3)\ \mathrm{s}^{-1}\,\mathrm{gauss}^{-1}\,.$

(v) The determination of the *faraday* by electrolysis; although more accurate values may soon be available we are at the moment confronted with two somewhat discrepant measurements on silver and on iodine. The iodine value is

(31) $F = Ne/c = (9\,652.15 + 0.13)\ \mathrm{emu/mole}$ (physical scale),

and the silver is

(32) $F = Ne/c = (9\,651.29 \pm 0.19)\ \mathrm{emu/mole}$ (physical scale) .

(vi) The *magnetic moment of the proton* in terms of the nuclear magneton. BLOCH and JEFFRIES obtain the value (uncorrected for diamagnetism)

(33) $\mu' = 2.792\,36 \pm 0.000\,10$

and HIPPLE, SOMMER and THOMAS obtain for the same quantity

(34) $\mu' = 2.792\,68 \pm 0.000\,03\,.$

(vii) The *short wavelength limit* of the continuous X-ray spectrum is unfortunately not known experimentally with sufficient accuracy to carry much weight in a least squares fitting. The difficulties here have been previously discussed [14]. It appears that the best experimental value is

(35) $hc^2/e\varLambda = (12370.8 + 1.0)\ \mathrm{emu}\,.$

In formulating the data on which our input equations of observation are to be based, care had been taken to do so in such a way as to avoid hidden correlations between the equations of observation. For example, the measurements resulting from observed values of crystal densities and their X-ray grating constants are almost invariably quoted as measurements of Avogadro's

number, N. To obtain N from these measurements, however, it is necessary to combine the results whith the cube of the conversion constant, Λ. Consequently we have equated the numeric which was actually measured to the appropriate function, $N\Lambda^3$, of our primary unknown. In other words we have been careful to see that no single measured quantity shall be involved as a substantial contributor to the error measures to two or more of the observational equations at once.

It will not be possible here to discuss the numerical details of the actual least squares analysis. These may be found elsewhere in the literature [15, 16]. Only the results and their evaluation will be presented. It has already been mentioned, probably more than once, that the experimental data are not entirely self-consistent. Thus, the two determinations of the Faraday constants (by iodine and silver voltameters) differ by 89 parts per million although each measurement claims to have an accuracy of 20 parts per million or better. The measurement of the proton magnetic moment by BLOCH and JEFFRIES is 115 parts per million higher than the measurement by HIPPLE, SOMMER and THOMAS. This is some three times larger than the standard deviation of the data would imply. The recently published reanalysis of this experiment [17] has resulted in a reduction of BLOCH and JEFFRIES' result by 109 parts per million, thus bringing it in excellent accord with HIPPLE, SOMMER and THOMAS. Furthermore, the modification of the inverse cyclotron by COLLINGTON, DELLIS, SANDERS and TURBERFIELD [18] yields a value which differs from TRIGGER's corrected BLOCH and JEFFRIES result by 21 ± 39 parts per million, and from HIPPLE, SOMMER and THOMAS omegatron value by 16 ± 18 parts per million. Professor DuMOND has already fully described the problems associated with the short wavelength limit.

A preliminary least squares analysis was carried out in 1952 with the data then available. This comprised thirteen equations in five variables, α, e, N, Λ and c. The value χ^2 (which is the minimum value of the function Q — the function whose minimization it is our object to obtain) was 52.1. This is a disturbingly large figure since there are only eight degrees of freedom and hence we would expect χ^2 to have the mean value 8 and to lie between the limits 2.73 and 15.51 with 90% confidence. The probability is less than one in a thousand that χ^2 would be larger than 26.1 entirely as a result of random error. We may say with almost complete certainty that there is a systematic error either in the observational data themselves or in the assignment of weights to those measurements.

The most convenient method for investigating the existence of systematic error in a set of data such as the one at hand is the comparison of the direct and indirect values of each input numeric as has been described earlier. Such a comparison was made and the implications are direct and conclusive [13]. The iodine value of the faraday for example is much more consistent with

the indirect value than is the silver faraday. The difference between the iodine value and the indirect value is 53 ± 26 parts per million, while the silver faraday differs from the indirect value by 142 ± 30 parts per million. The existence of sources of experimental systematic error in this measurement has been pointed out and this physical evidence strengthens the significance of the statistical implication.

Similarly the indirect value of the proton magnetic moment agrees with HIPPLE, SOMMER and THOMAS' value, thus indicating the possible existence of an error in the original inverse cyclotron work. This, of course, was later confirmed by TRIGGER's reanalysis [17] and the modified experiment carried out at Oxford by SANDERS et al. [18].

The determinations of the short wavelength limit of the continuous X-ray spectrum are all in disagreement with the indirect value. Hence we may tend to doubt the direct data; there are certainly valid experimental reasons why we should, and one is tempted to reject all of these data.

A word of caution is in order at this point. Of the thirteen items which were included in this preliminary analysis we have deduced reasons to reject five. This represents a rather drastic censoring and should be carried out with care. As soon as we eliminate one item the values of all of the other adjusted output values change. Hence strictly we should re-evaluate the least squares solution after each rejection in order to determine a new basis for the subsequent rejections. Furthermore, the decision to reject a datum as discrepant may very well depend upon the order in which other data are rejected. It is quite possible, in general, that a particular experimental result is in disagreement with the indirectly determined value, not because of an error in the experiment itself, but because some other error-ridden experiment is strongly distorting the indirect value. If the latter experiment were rejected first, the former one might well be retained. This is particularly liable to occur if the total degrees of freedom of the system is small so that each experimental item can contribute significantly to the overall result.

A completely different approach to the problem of determining which equations of the thirteen are likely to suffer from systematic error is afforded by an analysis of variance approach. This analysis involves the calculation of χ^2 for every possible subset of the thirteen equations and the comparison of these values of χ^2 to determine whether the value of 52.1 can be ascribed to the effect of one or a few items or whether all of the data contribute equally to this large value. In the first instance we would be willing to say that those items responsible for anomalously large contributions to χ^2 were, in fact, suffering from systematic error; in the second instance we would ascribe an increased standard error to all of the data and admit that if systematic errors existed they were so wide-spread as to defy unique identification.

A partial survey, although not complete, is adequate to confirm the implications of Table II: Silver coulometer value for the faraday constant and Bloch and Jeffries proton magnetic moment determination are in error and all of the short wavelength limit measurements are somewhat suspect. Actually suspicion does not fall strongly on the Bearden and Schwarz value but this is due more to its low weight than to its close agreement. Thus although its residue is in most comparisons as large or larger than the residues of the other

TABLE II. – *Direct and indirect evaluation of fundamental constants (1952 Adjustment).*

Constant		Direct value (experimental)	Indirect value	Least squares solution	Residue of least squares solution (ppm)
A		$1.002\,020 \pm 0.000\,030$	$1.002\,073 \pm 0.000\,015$	$1.0020\,63 \pm 0.000\,013$	43 ± 27
NA^3		$6\,061.79 \pm 0.23 \cdot 10^{20}$	$6\,062.90 \pm 0.39$	$6\,062.08 \pm 0.20$	48 ± 19
γ		$(26\,752.3 \pm 0.6)\ \mathrm{emu^{-1}\,s^{-1}}$	$26\,751.9 \pm 0.4$	$26\,752.0 \pm 0.5$	-11 ± 20
F	(I)	$(9\,652.15 \pm$ $\pm\,0.13)\ \mathrm{emu/mole}$	$9\,651.90 \pm 0.15$	$9\,652.01 \pm 0.10$	-11 ± 9
	(Ag)	$(9\,651.29 \pm$ $\pm\,0.19)\ \mathrm{emu/mole}$	$9\,652.21 \pm 0.11$	$9\,652.01 \pm 0.10$	78 ± 17
μ'	(BJ)	2.79237 ± 0.00010	2.79270 ± 0.00005	2.79267 ± 0.00003	109 ± 34
	(HST)	2.79268 ± 0.00003	2.79263 ± 0.00005	2.79267 ± 0.00003	-5 ± 6
X-ray S.W.L.	(FHD)	$(12\,370.02 \pm 0.63)\ \mathrm{emu\,cm}$	$12\,372.40 \pm 0.17$	$12\,372.23 \pm 0.16$	179 ± 48
	(BJW)	$(12\,371.03 \pm 0.48)\ \mathrm{emu\,cm}$	$12\,372.37 \pm 0.17$	$12\,372.23 \pm 0.16$	97 ± 39
	(BS)	$(12\,370.77 \pm 1.03)\ \mathrm{emu\,cm}$	$12\,372.28 \pm 0.17$	$12\,372.23 \pm 0.16$	119 ± 81

Remarks. – BJ = F. BLOCH and C. D. JEFFRIES; HST = J. A. HIPPLE, H. SUMMER and H. A. THOMAS; FHD = G. F. FELT, J. N. HARRIS and J. W. M. DuMOND; BJW = J. A. BEARDEN, F. T. JOHNSON and H. M. WATTS; BS = J. A. BEARDEN and G. SCHWARZ.

X-ray S.W.L. = Short wavelength limit of the continuous X-ray spectrum.

The column marked « Direct value » gives the experimental datum measured in the experiment which is indicated in the first column. The « Indirect value » is the one which would be deduced from a least squares analysis if the given experiment were excluded from consideration. If the same experiment is repeated by different observers only the single experiment in question (and not the others of the same kind) is excluded. The column marked « Residue of least squares solution » is the adjusted value of the 1952 analysis. The last column gives, in parts per million, the difference between the least squares solution and the input datum. The error quoted is the standard error of that difference.

[The fine structure splitting, ΔE_{d}, and the velocity of light have been omitted from this table because the direct measurements of these items have so much more weight than the indirect values that the elimination of numerical accuracy in the calculation precludes a meaningful evaluation of the latter].

short wavelength limit determinations, its contribution to χ^2 is much smaller because of the much larger standard error assigned to it.

The 1955 evaluation of the atomic constant [16] therefore rejected the silver faraday and BLOCH and JEFFRIES proton magnetic moment measurement. In addition the short wavelength limit values were replaced by a single expression based on the average of all the low voltage data as has been previously mentioned. It was also considered adequate to consider the velocity of light to be an auxiliary constant, not subject to least squares analysis. This rather drastic censoring of the data reduces the system to seven equations in four unknowns. The calculated value of χ^2 is then found to be 3.25 which is deceptively close to the expectation value, since the 90% confidence limits are $0.35 < \chi^2 < 7.81$. It can also be argued that the short wavelength limit measurements should be dropped entirely (and a strict adherence to our selection conditions would require this). In this case we decrease our system to one with only two degrees of freedom. The value of χ^2 is then only 0.44. This is at first sight quite low compared with the expectation value of 2.00; however the 90% confidence limits are 0.10 and 5.99. In fact, the probability that χ^2 might be lower than 0.44 is almost 20%, and this probability is further increased if we recognize that we have rather drastically censored all those data which contribute to large values of χ^2. Furthermore, the actual solution is not greatly changed by rejecting the short wavelength limit data entirely; the largest change in any output value is less than one fourth of the assigned standard deviation.

The system of seven equations in four unknowns which is obtained when we exclude those data which appear to be discordant with the remainder of our data is given below. As has been mentioned above we assume the velocity of light to be an exact constant in this analysis, its value being

$$(36) \qquad\qquad c = 299\,793.0 \text{ km/s} .$$

(The actual standard deviation to be assigned to this number is $\pm\, 0.3$ km/s, but this error has a negligible effect on the numerical results of the least squares analysis.)

The origin values and the linearized variables are taken to be:

$$(37) \qquad\qquad x_1 = \frac{\alpha - \alpha_0}{\alpha_0} \cdot 10^5 \quad \alpha_0 = 0.007\,297\,000 ,$$

$$(38) \qquad\qquad x_2 = \frac{e - e_0}{e_0} \cdot 10^5 \quad e_0 = 4.802\,200 \cdot 10^{-10} \text{ esu} ,$$

$$(39) \qquad\qquad x_3 = \frac{N - N_0}{N_0} \cdot 10^5 \quad N_0 = 0.602\,500\,0 \cdot 10^{24} \text{ mole}^{-1},$$

$$(40) \qquad\qquad x_4 = \frac{\Lambda - \Lambda_0}{\Lambda_0} \cdot 10^5 \quad \Lambda_0 = 1.002\,020\,0 .$$

With these definitions the observational equations become:

			Weight	Experimental source
(41)	$x_4 =$	0.0	0.11,	$\Lambda = 1.002\ 02 \pm 0.000\ 03$
(42)	$x_3 + 3x_4 =$	3.5	0.07,	$N\Lambda^3$ (Birge's average)
(43)	$x_1 \qquad\quad =$	4.0	4.92,	Fine structure splitting in deuterium
(44)	$3x_1 - x_2 \qquad = -$	2.3	0.19,	Gyromagnetic ratio of proton
(45)	$x_2 + x_3 \quad =$	11.1	0.58,	Iodine faraday, electrochemistry
(46)	$-3x_1 + 2x_2 + x_3 \quad =$	13.5	0.83,	Magnetic moment of proton (omegatron)
(47)	$-x_1 + x_2 \quad - x_4 = -$	5.6	0.015,	Short wavelength limit (low voltage)

The normal equations are formed according to the usual rules; they are

$$
(48)\quad
\begin{cases}
(\alpha) & 14.115x_1 & - 5.565x_2 & - 2.490x_3 & + 0.015x_4 = & - 15.162 \\
(e) & - 5.565x_1 & + 4.105x_2 & + 2.240x_3 & - 0.015x_4 & 29.201 \\
(N) & - 2.490x_1 & + 2.240x_2 & + 1.480x_3 & + 0.210x_4 = & 17.888 \\
(\Lambda) & 0.015x_1 & - 0.015x_2 & + 0.210x_3 & + 0.755x_4 = & 0.819
\end{cases}
$$

It should be noted here that the normal equations are to be formed directly from the coefficients of the quadratic form, Q, as given in equation (16) and that no simplification or cancellation should be made in the equations. This is because the coefficients of the normal equations have a significance which is not limited to the determination of the values of the x_i. An alteration in the equations which would not alter the solution (such as the cancellation of a common factor in one of the equations, or even the re-ordering of the equations in the set) can destroy the identifications of the coefficients of the normal equations as the weight matrix of the solution.

The error (or variance) matrix of the solution is the inverse of the matrix of the coefficients of the normal equations. Hence the error matrix is

$$
(49)\quad V =
\begin{pmatrix}
0.198\ 9 & - 0.5760 & 0.560\ 3 & 0.163\ 3 \\
0.576\ 0 & 3.4478 & - 4.431\ 9 & 1.289\ 8 \\
- 0.560\ 3 & - 4.4319 & 6.716\ 7 & - 1.945\ 2 \\
0.163\ 3 & 1.2898 & - 1.945\ 2 & 1.887\ 9
\end{pmatrix}
\begin{matrix} \alpha \\ e \\ N \\ \Lambda \end{matrix}
$$
$$
\begin{matrix} \alpha & \quad e & \quad N & \quad \Lambda \end{matrix}
$$

and the solution for the variables is

(50)
$$\begin{cases} x_1 = 3.92 \\ x_2 = 13.72 \\ x_3 = -2.37 \\ x_4 = 1.94 \end{cases}$$

The value of χ^2 for this solution is

(51) $\chi^2 = 3.25$.

This is to be compared to the value 52.1 which was obtained in the preliminary adjustment. The major change in the variables caused by the deletion of those data suspected of systematic error is an increase in x_3 of 2.23 and an attendant decrease in x_4 of 2.35; this corresponds to a change in Avogadro's number N of 22.3 ppm and a change in the conversion factor Λ of -23.5 ppm.

The value of χ^2 can be expressed equivalently in terms of the ratio of external to internal consistency

$$r_e/r_i = \sqrt{\chi^2/(n-q)} = 1.041 .$$

The 90% confidence interval for r_e/r_i with 3 degrees of freedom is $0.594 < r_e/r_i < 2.795$. Hence there is no particular compulsion to use internal rather than external consistency measures; our data tell us only that the two measures are consistent. We shall, however, quote our final error matrix in terms of external consistency since this is the larger measure (albeit not significantly so). Furthermore, the error matrix will be re-expressed in terms of relative parts per million in the primary variables.

The error matrix and correlation coefficients are given in Tables III and IV.

TABLE III. – *Covariance matrix (1955 Adjustment).*
Elements of the matrix are in units of (relative parts per million)^2.

	e	m	h	α	Λ	N	F
e	374	560	685	62	140	-480	-107
m	560	940	1 057	60	226	-778	-218
h	685	1 057	1 246	103	262	-899	-216
α	62	60	103	22	18	-61	2
Λ	140	226	262	18	204	-211	-71
N	-480	-778	-899	-61	-211	726	246
F	-107	-218	-216	2	-71	246	141

TABLE IV. – *Correlation coefficients (1955 Adjustment)*.

	e	m	h	α	A	N	F
e	1.00	0.95	0.99	·0.70	0.51	— 0.92	— 0.47
m	0.95	1.00	0.97	0.43	0.52	— 0.94	— 0.60
h	0.99	0.97	1.00	0.63	0.52	— 0.94	— 0.51
α	0.70	0.43	0.63	1.00	0.27	— 0.48	0.03
A	0.51	0.52	0.52	0.27	1.00	— 0.55	— 0.42
N	— 0.92	— 0.94	— 0.94	— 0.48	— 0.55	1.00	0.77
F	— 0.47	— 0.60	— 0.51	0.03	— 0.42	0.77	1.00

The error matrix of Table III has been extended to include the variables m, h and F. These variables were eliminated from the least squares solution by the use of auxiliary equations. It is, however, inconvenient to be forced to use these equations in order to express m, h or F in terms of our primary variables each time one needs to compute correlation coefficients between these variables and other variables. The variances and correlation coefficients for these secondary variables can be calculated in terms of the variance matrix of the primary variables. Thus, although we have produced a matrix of seventh order it is still only of the fourth rank since three of the rows (or columns) can be expressed as linear combinations of the other rows (or columns). The near equality of m and h columns of the matrix of correlation coefficients (Table IV) is a reflection of the strong coupling between the adjusted values of these two variables.

This is a result of the relationship

$$m = \frac{2R_\infty h}{c\alpha^2},$$

and the fact that the relative standard error of α is much smaller than the error in either m or h.

7. – Output values. Recommended (1955) least squares adjustment.

The 1955 adjustment is based, as has been explained earlier, on the determination of four primary unknowns α, e, N and A, by least-squares adjustment from seven independent sources of experimental input data combined with a number of auxiliary constants whose values are known so much more accurately than the aforementioned input data that they are negligible error contributors.

The physical scale of atomic weights is used almost exclusively in these tables, the Sackur-Tetrode constant (given on both physical and chemical scales) being the sole exception. The conversion factor between these scales, $r = 1.000\,272 + 0.000\,005$ calculated by R. T. BIRGE on the basis of the abun-

dance ratio for the oxygen isotopes, $^{16}O : ^{18}O : ^{17}O = (506 \pm 10):1:(0.204 \pm 0.008)$ is adopted here as a definition for the chemical scale. These abundance ratios, and the value or r implied by them, are subject to variation depending on the source of the oxygen. The value of the $^{16}O/^{18}O$ ratio can vary from approximately 495 for oxygen from air or carbonates to 515 for oxygen from water and rocks. Corresponding to this variation is a variation in the value of r from 1.002 78 to 1.002 68. The International Commission on Atomic Weights is at present (1956) considering the arbitrary redefinition of the chemical scale of atomic weights in terms of the physical scale and the value $r = 1.002\,75$. The accuracy ascribed to our adopted value is such that these two numbers do not differ significantly (approximately one half the standard error).

The new Kelvin scale of temperature adopted October 1954 in Paris at the Tenth General Conference on Weights and Measures is here used. On this scale the triple point of water is assigned the temperature 273.16 °K exactly. This changes the numerical value of the gas constant, R_0, slightly from that used in earlier evaluations and gives the value of the ice point as (273.150 0 $\pm$ 0.000 2) °K. Absolute electrical units are used exclusively, the « international » electrical units having been abolished in 1948.

TABLE V. – *Auxiliary constants.*

Rydberg wave number for infinite mass
$$R_\infty = (109\,737.309 \pm 0.012)\ \text{cm}^{-1}$$

Rydberg wave numbers for the light nuclei
$$R_H = (109\,677.576 \pm 0.012)\ \text{cm}^{-1}$$
$$R_D = (109\,707.419 \pm 0.012)\ \text{cm}^{-1}$$
$$R_{^3He} = (109\,717.345 \pm 0.012)\ \text{cm}^{-1}$$
$$R_{^4He} = (109\,722.267 \pm 0.012)\ \text{cm}^{-1}$$

Velocity of light
$$c = (299\,793.0 \pm 0.3)\ \text{km s}^{-1}$$

Atomic mass of neutron
$$M_n = 1.008982 \pm 0.000003 \ \text{(physical scale)}$$

Atomic mass of hydrogen
$$H = 1.008142 \pm 0.000003 \ \text{(physical scale)}$$

Atomic mass ratio of hydrogen to proton
$$H/M_p = 1.00054461 \quad \text{(computed using atomic mass of electron}$$
$$Nm = 0.00054875)\ \text{(physical scale)}$$

Atomic mass of the proton
$$M_p = 1.007593 \pm 0.000003 \ \text{(physical scale)}$$

Atomic mass of deuterium
$$D = 2.014735 \pm 0.000006 \ \text{(physical scale)}$$

TABLE V: *continued.*

Atomic mass ratio of deuterium to deuteron
$$D/M_d = 1.000\,272\,44 \text{ (computed using atomic mass of electron,}$$
$$Nm = 0.000\,548\,75) \text{ (physical scale))}.$$

Ratio of electron magnetic moment to proton magnetic moment without diamagnetic correction
$$[M_p/(Nm\mu')](1 + \alpha/2\pi - 2.973\alpha^2/\pi^2) = 658.228\,8 + 0.000\,4$$

Anomalous magnetic moment of electron
$$\mu_e/\mu_0 = (1 + \alpha/2\pi - 2.973\alpha^2/\pi^2) = 1.001\,145\,36$$
$$\text{(computed using the value } 1/\alpha = 137.037)$$

Gas constant per mole
$$R_0 = (8.316\,96 \pm 0.000\,34) \cdot 10^7 \text{ erg mole}^{-1} \text{ deg}^{-1} \text{ (physical scale)}$$

Standard volume of a perfect gas
$$V_0 = (22\,420.7 \pm 0.6) \text{ cm}^3 \text{ atm mole}^{-1} \text{ (physical scale)}$$

TABLE VI. – *Least squares adjusted output values, 1955.*

The quantity following each $\pm$ sign is the standard error. Attention is called to the fact that the quantities in this Table are observationally correlated so that in the computation of the error measures of derived values dependent on two or more of the values in this Table the error matrix of Table III must be used.

Avogadro's constant
$$N = (6.024\,86 \pm 0.000\,16) \cdot 10^{23} \text{ mole}^{-1} \text{ (physical scale)}$$

Loschmidt's constant
$$L_0 = N/V_0 = (2.687\,19 \pm 0.000\,10) \cdot 10^{19} \text{ cm}^{-3} \text{ (physical scale)}$$

Electronic charge
$$e = (4.802\,86 \pm 0.000\,09) \cdot 10^{-10} \text{ esu}$$
$$e' = e/c = (1.602\,06 \pm 0.000\,03) \cdot 10^{-20} \text{ emu}$$

Electron rest mass
$$m = (9.108\,3 \pm 0.000\,3) \cdot 10^{-28} \text{ g}$$

Proton rest mass
$$m_p = M_p/N = (1.672\,39 \pm 0.000\,04) \cdot 10^{24} \text{ g}$$

Neutron rest mass
$$m_n = M_n/N = (1.674\,70 \pm 0.000\,04) \cdot 10^{-24} \text{ g}$$

Planck's constant
$$h = (6.625\,17 \pm 0.000\,23) \cdot 10^{-27} \text{ erg s}$$
$$h/2\pi = (1.054\,43 \pm 0.000\,04) \cdot 10^{-27} \text{ erg s}$$

Conversion factor from Siegbahn X-units to milliangstroms
$$\Lambda = \lambda_g/\lambda_s = 1.002\,039 \pm 0.000\,014$$

TABLE VI: *continued.*

Faraday constant

$$F = Ne = (2.893\,66 \pm 0.000\,03) \cdot 10^{14} \text{ esu mole}^{-1}$$
$$F' = Ne/c = (9\,652.19 \pm 0.11) \text{ emu mole}^{-1} \quad \text{(physical scale)}$$

Charge-to-mass ratio of the electron

$$e/m = (5.273\,05 \pm 0.000\,07) \cdot 10^{17} \text{ esu g}^{-1}$$
$$e/mc = (1.758\,90 \pm 0.000\,02) \cdot 10^{7} \text{ emu g}^{-1}$$

Ratio h/e

$$h/e = (1.379\,42 \pm 0.000\,02) \cdot 10^{-17} \text{ erg s esu}^{-1}$$

Fine structure constant

$$\alpha = e^2/\hbar c = (7.297\,29 \pm 0.000\,03) \cdot 10^{-3}$$
$$1/\alpha = 137.037\,3 \pm 0.000\,6$$
$$\alpha/2\pi = (1.161\,398 \pm 0.000\,005) \cdot 10^{-3}$$
$$\alpha^2 = (5.325\,04 \pm 0.000\,05) \cdot 10^{-5}$$
$$1 - (1 - \alpha^2)^{\frac{1}{2}} = (0.266\,252 \pm 0.000\,002) \cdot 10^{-4}$$

Atomic mass of the electron

$$Nm = (5.487\,63 \pm 0.000\,06) \cdot 10^{-4} \quad \text{(physical scale)}$$

Ratio of mass of hydrogen to mass of proton (*)

$$\mathrm{H}/M_\mathrm{p} = \left[1 - \frac{Nm}{\mathrm{H}}\left(1 - \frac{1}{2}\alpha^2\right)\right]^{-1} =$$
$$= 1.000\,544\,613 \pm 0.000\,000\,006$$

Atomic mass of proton

$$M_\mathrm{p} = \mathrm{H} - Nm = 1.007593 \pm 0.000003 \quad \text{(physical scale)}$$

Ratio proton mass to electron mass

$$M_\mathrm{p}/(Nm) = 1\,836.12 \pm 0.02$$

Reduced mass of electron in hydrogen atom

$$\mu = mM_\mathrm{p}/\mathrm{H} = (9.103\,4 \pm 0.000\,3) \cdot 10^{-28} \text{ g}$$

Schrödinger constant for a fixed nucleus

$$2m/\hbar^2 = (1.638\,36 \pm 0.000\,07) \cdot 10^{27} \text{ erg}^{-1} \text{ cm}^{-2}$$

Schrödinger constant for the hydrogen atom

$$2\mu/\hbar^2 = (1.637\,48 \pm 0.000\,07) \cdot 10^{27} \text{ erg}^{-1} \text{ cm}^{-2}$$

First Bohr radius

$$a_0 = \hbar^2/(me^2) = \alpha/(4\pi R_\infty) =$$
$$= (5.291\,72 \pm 0.000\,02) \cdot 10^{-9} \text{ cm}$$

Radius of electron orbit in normal ^{1}H, referred to center of mass

$$a_0' = a_0(1 - \alpha^2)^{\frac{1}{2}} = (5.291\,58 \pm 0.000\,02) \cdot 10^{-9} \text{ cm}$$

Separation of proton and electron in normal ^{1}H

$$a_0'' = a_0' R_\infty/R_\mathrm{H} = (5.294\,46 \pm 0.000\,02) \cdot 10^{-9} \text{ cm}$$

(*) The binding energy of the electron in the hydrogen atom has been included in the quantity. The mass of the electron when found in the hydrogen atom is not m, but more correctly $m(1 - 1/2\alpha^2 + ...)$.

TABLE VI: *continued*.

Compton wavelength of the electron

$$\lambda_{C.,e} = h/(mc) = \alpha^2/2R_\infty = (24.262\ 6 \pm 0.000\ 2)\cdot 10^{-11}\ \text{cm}$$
$$\lambdabar_{C.,e} = \lambda_{C.,e}/2\pi = (3.861\ 51 \pm 0.000\ 04)\cdot 10^{-11}\ \text{cm}$$

Compton wavelength of the proton

$$\lambda_{C.,p} = h/m_p c = (13.214\ 1 \pm 0.000\ 2)\cdot 10^{-14}\ \text{cm}$$
$$\lambdabar_{C.,p} = \lambda_{C.,p}/2\pi = (2.103\ 08 \pm 0.000\ 03)\cdot 10^{-14}\ \text{cm}$$

Compton wavelength of the neutron

$$\lambda_{C.,n} = h/m_n c = (13.195\ 9 \pm 0.000\ 2)\cdot 10^{-14}\ \text{cm}$$
$$\lambdabar_{C.,n} = \lambda_{C.,n}/2\pi = (2.100\ 19 \pm 0.000\ 03)\cdot 10^{-14}\ \text{cm}$$

Classical electron radius

$$r_0 = e^2/(mc^2) = \alpha^3/(4\pi R_\infty) =$$
$$= (2.817\ 85 \pm 0.000\ 04)\cdot 10^{-13}\ \text{cm}$$
$$r_0^2 = (7.940\ 30 \pm 0.000\ 21)\cdot 10^{-26}\ \text{cm}$$

Thomson cross section

$$(8/3)\pi r_0^2 = (6.652\ 05 \pm 0.000\ 18)\cdot 10^{-25}\ \text{cm}^2$$

Fine structure doublet separation in hydrogen

$$\Delta E_H = (1/16)R_H \alpha^2[1 + \alpha/\pi + (5/8 - 5.946/\pi^2)\alpha^2] =$$
$$= (0.365\ 871 \pm 0.000\ 003)\ \text{cm}^{-1}$$
$$= (10\ 968.56 \pm 0.10)\ \text{Hz}^{-1}$$

Fine structure separation in deuterium

$$\Delta E_D = \Delta E_H R_D/R_H = (0.365\ 970 \pm 0.000\ 003)\ \text{cm}^{-1}$$
$$= (10\ 971.54 \pm 0.10)\ \text{Hz}^{-1}$$

Zeeman displacement per gauss

$$e/4\pi mc^2 = (4.668\ 85 \pm 0.000\ 06)\cdot 10^{-5}\ \text{cm}^{-1}\ \text{G}^{-1}$$

Boltzmann's constant

$$k = R_0/N = (1.380\ 44 \pm 0.000\ 07)\cdot 10^{-16}\ \text{erg deg}^{-1}$$
$$= (8.616\ 7 \pm 0.000\ 4)\cdot 10^{-5}\ \text{eV deg}^{-1}$$
$$1/k = (11\ 605.4 \pm 0.5)\ \text{deg eV}^{-1}$$

First radiation constant

$$c_1 = 8\pi hc = (4.991\ 8 \pm 0.000\ 2)\cdot 10^{-15}\ \text{erg cm}$$

Second radiation constant

$$c_2 = hc/k = (1.438\ 80 \pm 0.000\ 07)\ \text{cm deg}$$

Atomic specific heat constant

$$c_2/c = h/k = (4.799\ 31 \pm 0.000\ 23)\cdot 10^{-11}\ \text{s deg}$$

Wien displacement law constant (*)

$$\lambda_{max}T = c_2/4.965\ 11\ 23 = (0.289\ 782 \pm 0.000\ 013)\ \text{cm deg}$$

Stefan-Boltzmann constant

$$\sigma = (\pi^2/60)(k^4/\hbar^3 c^2) = (0.566\ 87 \pm 0.000\ 10)\cdot 10^{-4}\ \text{erg cm}^{-2}\ \text{deg}^{-4}\ \text{s}^{-1}$$

(*) The numerical constant 4.965 114 23 is the root of the transcendental equation
$x = 5(1 - \exp[-x])$.

TABLE VI: *continued.*

Sackur-Tetrode constant
$$S_0/R_0 = \tfrac{5}{2} + \ln\left[(2\pi R_0)^{\frac{3}{2}}h^{-3}N^{-4}\right] =$$
$$= -5.57324 \pm 0.00007$$
$$S_0 = -(46.3524 \pm 0.0020) \cdot 10 \text{ erg mole}^{-1}\text{deg}^{-1}$$
<div align="right">(physical scale)</div>

Sackur-Tetrode constant
$$S_0/R_{0\mathrm{Ch}} = -5.57256 \pm 0.00007$$
$$S_0 = -(46.3467 \pm 0.0020) \cdot 10^7 \text{ erg mole}^{-1}\text{ deg}^{-1}$$
<div align="right">(chemical scale)</div>

Bohr magneton
$$\mu_0 = he/4\pi mc = \tfrac{1}{2}e\hbar_{\mathrm{C.,e}}^2$$
$$= (0.92731 \pm 0.00002) \cdot 10^{-20} \text{ erg G}^{-1}$$

Anomalous electron moment correction
$$1 + \alpha/2\pi - 2.973\alpha^2/\pi^2 = \mu_{\mathrm{e}}/\mu_0 = 1.001145358 \pm 0.000000005$$
<div align="center">(computed using adjusted value $1/\alpha = 137.0373 \pm 0.6000$)</div>

Magnetic moment of the electron
$$\mu_{\mathrm{c}} = (0.92837 \pm 0.00002) \cdot 10^{-20} \text{ erg G}^{-1}$$

Nuclear magneton
$$\mu_{\mathrm{n}} = he/(4\pi m_{\mathrm{p}}c) = \mu_0 Nm/\mathrm{H}^+ =$$
$$= (0.505038 \pm 0.000018) \cdot 10^{-23} \text{ erg G}^{-1}$$

Proton moment
$$\mu_{\mathrm{p}} = (2.79275 \pm 0.00003) \text{ nuclear magnetons}$$
$$= (1.41044 \pm 0.00004) \cdot 10^{-23} \text{ erg G}^{-1}$$

Gyromagnetic ratio of the proton in hydrogen (uncorrected for diamagnetism)
$$\gamma' = (2.67523 \pm 0.00004) \cdot 10^4 \text{ rad s}^{-1} \text{ G}^{-1}$$

Gyromagnetic ratio of the proton (corrected)
$$\gamma = (2.67530 \pm 0.00004) \cdot 10^4 \text{ rad s}^{-1} \text{ G}^{-1}$$

Multiplier of (Curie constant)$^{\frac{1}{2}}$ $x^{\frac{1}{2}}$, to give magnetic moment per molecule
$$(3k/N(x)^{\frac{1}{2}} = 2.6178 \pm 0.00010) \cdot 10^{-20} \text{ (erg mole deg}^{-1})^{\frac{1}{2}}$$

Mass-energy conversion factors
$$1 \text{ g} = (5.61000 \pm 0.00011) \cdot 10^{26} \text{ MeV}$$
$$1 \text{ electron mass} = (0.510976 \pm 0.000007) \text{ MeV}$$
$$1 \text{ atomic mass unit} = (931.141 \pm 0.010) \text{ MeV}$$
$$1 \text{ proton mass} = (938.211 \pm 0.010) \text{ MeV}$$
$$1 \text{ neutron mass} = (939.505 \pm 0.010) \text{ MeV}$$

Quantum energy, E, conversion factors
$$1 \text{ eV} = (1.60206 \pm 0.00003) \cdot 10^{-12} \text{ erg}$$
$$E/\tilde{\nu} = hc = (1.98618 \pm 0.00007) \cdot 10^{-16} \text{ erg cm}$$
$$E\lambda_g = (12397.67 \pm 0.22) \cdot 10^{-8} \text{ eV cm}$$
$$E\lambda_s = (12372.44 \pm 0.16) \text{ kV } X\text{-units}$$
$$E/\nu = (6.62517 \pm 0.00023) \cdot 10^{-27} \text{ erg s}$$
$$= (4.13541 \pm 0.00007) \cdot 10^{-15} \text{ eV s}$$

TABLE VI: *continued.*

$$\tilde{\nu}/E = (5.034\,79 \pm 0.000\,17) \cdot 10^{15} \text{ cm}^{-1} \text{ erg}^{-1}$$
$$= (8\,066.03 \pm 0.14) \text{ cm}^{-1} \text{ eV}^{-1}$$
$$\nu/E = (1.509\,40 \pm 0.000\,05) \cdot 10^{26} \text{ s}^{-1} \text{ erg}^{-1}$$
$$= (2.418\,14 \pm 0.000\,04) \cdot 10^{14} \text{ s}^{-1} \text{ eV}^{-1}$$

de Broglie wavelengths $\lambda_{\text{B.,}}$ of elementary particles (*)

Electrons

$$\lambda_{\text{B.,e}} = (7.273\,77 \pm 0.000\,06) \text{cm}^2 \text{ s}^{-1}/\text{V}$$
$$= (1.552\,257 \pm 0.000\,016) \cdot 10^{-13} \text{ cm } (\text{erg})^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$
$$= (1.226\,378 \pm 0.000\,010) \cdot 10^{-7} \text{ cm } (\text{eV})^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$

Protons

$$\lambda_{\text{B.,p}} = (3.961\,49 \pm 0.000\,05) \cdot 10^{-3} \text{ cm}^2 \text{ s}^{-1}/\text{V}$$
$$= (3.622\,53 \pm 0.000\,08) \cdot 10^{-15} \text{ cm } (\text{erg})^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$
$$= (2.862\,02 \pm 0.000\,04) \cdot 10^{-9} \text{ cm } (\text{eV})^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$

Neutrons

$$\lambda_{\text{B.,n}} = (3.956\,03 \pm 0.000\,05) \cdot 10^{-3} \text{ cm}^2 \text{ s}^{-1}/\text{V}$$
$$= (3.602\,04 \pm 0.000\,08) \cdot 10^{-15} \text{ cm } (\text{erg}^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$
$$= (2.860\,05 \pm 0.000\,04) \cdot 10^{-9} \text{ cm } (\text{eV})^{\frac{1}{2}}/(E)^{\frac{1}{2}}$$

Energy of 2 200 m/s neutron

$$E_{2200} = (0.025\,297\,3 \pm 0.000\,000\,3) \text{ eV}$$
$$T_{2200} = (293.585 \pm 0.012) \text{ °K}$$
$$= (20.435 \pm 0.012) \text{ °C}$$

The Rydberg and related derived constants

$$R_\infty = (109\,737.309 \pm 0.012) \text{ cm}^{-1}$$
$$R_\infty c = (3.289\,848 \pm 0.000\,003) \cdot 10^{15} \text{ s}^{-1}$$
$$R_\infty hc = (2.179\,58 \pm 0.000\,07) \cdot 10^{-11} \text{ erg}$$
$$Rhc^2 e^{-1} \cdot 10^{-8} = (13.604\,88 \pm 0.000\,22) \text{ eV}$$

Hydrogen ionization potential

$$I_0 = R_{\text{H}}(hc^2/e)(1 + \tfrac{1}{4}\alpha^2 + \ldots) \cdot 10^{-8} =$$
$$= (13.597\,65 \pm 0.000\,22) \text{ eV}$$

(*) These formulae apply only to nonrelativistic velocities. If the velocity of the particle is not negligible compared to the velocity of light, c, or the energy not negligible compared to the rest mass energy, we must use $\lambda_{\text{B.,D}} = \lambda_C[\varepsilon(\varepsilon+2)]^{-\frac{1}{2}}$ where λ_C is the appropriate Compton wavelenght of the particle in question and ε is the kinetic energy measured in units of the particle rest-mass.

REFERENCES

[1] J. W. M. DuMond and E. R. Cohen: *Rev. Mod. Phys.*, **25**, 691 (1953).
[2] E. R. Cohen: *Phys. Rev.*, **88**, 353 (1952).
[3] R. T. Birge: *Reports on Progress in Physics*, **8**, 90 (1941).

[4] W. N. BOND: *Phil. Mag.*, **10**, 994 (1930); **12**, 632 (1931).

[5] R. T. BIRGE: *Phys. Rev.*, **40**, 228 (1932).

[6] R. A. BETH: *Phys. Rev.*, **53**, 681 (1938).

[7] J. W. M. DuMOND: *Phys. Rev.*, **56**, 153 (1939).

[8] C. F. GAUSS: *Theoria Combinations Observationum Erroribus Minimis Obnoxiae, Werke*, Bd. **4** (Göttingen, 1873).

[9] E. WHITTAKER and G. ROBINSON: *Calculus of Observations* (London, 1944), p. 218.

[10] E. WHITTAKER and G. ROBINSON: *Calculus of Observations* (London, 1944), p. 224.

[11] E. R. COHEN: *Rev. Mod. Phys.*, **25**, 705 (1953).

[12] A. C. AITKEN: *Proc. Roy. Soc. Edinburgh*, **55**, 42 (1935).

[13] E. R. COHEN: *Phys. Rev.*, **101**, 1641 (1956).

[14] J. W. M. DuMOND: Preceding paper in this issue.

[15] J. W. M. DuMOND and E. R. COHEN: *Rev. Mod. Phys.*, **25**, 691 (1953).

[16] E. R. COHEN, J. W. M. DuMOND, T. W. LAYTON and J. S. ROLLETT: *Rev. Mod. Phys.*, **27**, 363 (1955).

[17] D. R. TRIGGER: *Bull. Am. Phys. Soc.*, **1**, 220 (Abstract UA8) (1956).

[18] D. J. COLLINGTON, A. N. DELLIS, J. H. SANDERS and K. C. TURBERFIELD: *Phys. Rev.*, **99**, 1622 (1955).

[19] R. A. FISHER: *Statistical Methods for Research Workers* (Edinburgh, 1952), Table XIV.