

A Critical Examination of the Statistical Method for the Analysis of Gas Electron Diffraction Data

BY YONEZO MORINO, KOZO KUCHITSU AND YOSHITADA MURATA

Department of Chemistry, Faculty of Science, The University of Tokyo, Bunkyo-ku, Tokyo, Japan

(Received 15 July 1963 and in revised form 4 April 1964)

In view of the importance of gas electron diffraction to the precise determination of molecular structure, a critical study is made of the method of least-squares analysis in order to attain as statistically accurate results as possible from experimental data.

It is pointed out that 'observations' on the molecular intensity curve $qM(q)$ and on the radial distribution curve $f(r)$ should be regarded as 'indirect', if the criterion of least squares is applied either to $qM(q)$ or to $f(r)$. A general formula for the weight function which is related to indirect observations is derived, and it is shown by using the formula that there is intensive interdependence among the nearest neighbors of $f(r)$.

A practical method is devised for estimating a reasonable weight function for $qM(q)$ from the differences between the observed and calculated values of $qM(q)$.

It is remarked that the effect of anharmonicity parameters κ on the internuclear distances r must be taken into account even for non-hydride molecules in order to claim high accuracy for the distances.

A method for estimating random errors by making use of the standard errors denoted as σ_1 (the internal consistency) and σ_2 (the reproducibility) is proposed, and the importance of σ_2 as a measure of random errors is pointed out.

The method mentioned above is demonstrated by taking an example, silicon tetrachloride, and the standard errors of the internuclear distances and the mean amplitudes are estimated.

The technique of gas electron diffraction has recently reached the stage where internuclear distances and mean vibrational amplitudes in polyatomic molecules can be determined with experimental errors of a few thousandths of 1 Å. Such precise information is of fundamental importance to structural chemistry, since various subtle features of molecular structure, *e.g.* the effect of molecular vibrations on the distances are shown to be of this order of magnitude. Accordingly, as extensively discussed at the International Symposium on Electron Diffraction in Kyoto (*Proc. Intern. Conf. Mag. Crystallogr.*, 1962), one must strive during the analysis of scattering intensities to derive as reliable information as possible with an estimated uncertainty which has full statistical significance.

Application of the least-squares method to this problem was initiated by Hamilton (1954), and later extended by Bastiansen, Hedberg & Hedberg (1957). The method was also applied to the radial distribution curve with a slight modification by Bonham & Bartell (1959). The present paper is concerned with a critical study of the analytical procedures currently in use in searching for a reasonable estimate of the most probable values and the standard errors of molecular parameters derived from experiment.

General principle of least squares

On the basic assumption that an ordinary theory of errors is applicable to the measurement of scattering

intensities, the principle of least squares is conveniently formulated in matrix notation as stated by Bastiansen, Hedberg & Hedberg (1957), Hedberg & Iwasaki (1962, 1964), and Morino, Kuchitsu, Iijima & Murata (1962). Let \mathbf{F} be a set of calculated values of the n observables expressed in an n row and one column matrix whose elements are defined by a set of linear combinations of m independent parameters; it follows that

$$\mathbf{F} = \mathbf{M} - \mathbf{V} = \mathbf{A}\mathbf{X}, \quad (1)$$

where \mathbf{M} , \mathbf{V} , \mathbf{A} , and \mathbf{X} are matrices of the observations, the residuals, the coefficients, and the estimates of the parameters, respectively. The condition of least squares requires that

$$\mathbf{V}^* \mathbf{P} \mathbf{V} = \text{minimum, or } \mathbf{A}^* \mathbf{P} \mathbf{V} = 0, \quad (2)$$

where \mathbf{P} is a weight matrix and is supposed to be symmetric. By solving the normal equations,

$$\mathbf{B}\mathbf{X} = \mathbf{A}^* \mathbf{P} \mathbf{M}, \quad \text{where } \mathbf{B} = \mathbf{A}^* \mathbf{P} \mathbf{A}, \quad (3)$$

unknown elements of \mathbf{X} are determined. The standard error σ_i of the element x_i of the estimates \mathbf{X} is given by

$$\sigma_i = (\mathbf{B}^{-1})_{ii}^{\frac{1}{2}} [\mathbf{V}^* \mathbf{P} \mathbf{V} / (n - m)]^{\frac{1}{2}}. \quad (4)$$

Since our problem deals with non-linear functions, we start with a set of trial parameters \mathbf{X}_0 . The solutions $\Delta\mathbf{X}$ of the normal equations are added to the initial \mathbf{X}_0 to obtain revised parameters, and the process is recycled until the parameters \mathbf{X} converge.

Although the procedure described above is well established in principle, several practical problems must be handled carefully in its application. The most important questions may be the following:

- How should the interval of observations and the weight matrix be taken?
- To what function should the observations be fitted?
- How should random errors be estimated?

Weight matrix

In any of the previous studies, the weight matrix was assumed to be diagonal, and the elements were assumed to form a uniform function of q (or r), which was introduced somewhat arbitrarily in conformance with general experience. A typical function with an exponential damping in large q is shown in Fig. 1(a).

It will be shown in the following, however, that the weight matrix for the radial distribution curve $f(r)$ may not be diagonal; nor may the weight for the molecular intensity $qM(q)$ be a uniform function of q , even if the weight matrix for our initial observations on the microphotometer trace is diagonal and uniform. This is because our observations on $qM(q)$ or on $f(r)$ must be regarded as 'indirect'. Since we usually apply the criterion of least squares (equation (2)) either to $qM(q)$ or to $f(r)$, both of which are derived after a series of manipulations from our original (direct) measurement on the photometer trace, equations (3) and (4) should be modified accordingly.

As explained in Appendix I, equations (1) through (4) are essentially unchanged for indirect observations provided that the weight matrix \mathbf{P} is replaced by a matrix Π which is given by

$$\Pi = (\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}, \quad (5)$$

where \mathbf{G} represents the matrix of derivatives which combines the direct (\mathbf{M}) and indirect (μ) observations by equation (A-1). If the dependence of μ on \mathbf{M} is known, \mathbf{G} can be calculated by taking the derivatives. Equation (5) is a useful relation for estimating the weight matrix to be used in our analysis.

Since the elements of \mathbf{G} for $qM(q)$ (non-square in this case) which refer to different q values are all zero (see Appendix II, equation (A-17)), the weight matrix Π_m for $qM(q)$ is diagonal if the original matrix \mathbf{P} in equation (5) is diagonal, and hence, if our observations on the photometer chart are all mutually independent. As far as this condition is satisfied, the decrease in the interval of observations, Δq , and hence the increase in the number of observations, will make the standard error σ_i in equation (4) smaller. Obviously, however, Δq should not be made smaller beyond a certain limit, since correlation among the observations will then become appreciable (Bastiansen, Hedberg & Hedberg, 1957). Off-diagonal elements, by which the correlation can be dealt with, must therefore be introduced in

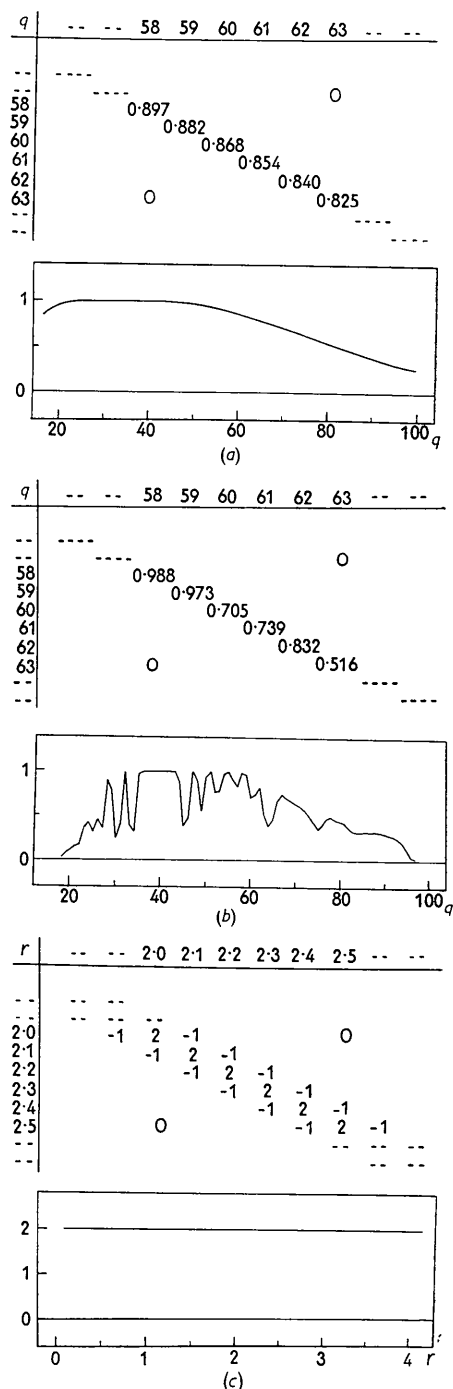


Fig. 1. Typical weight matrices (a, b) for $qM(q)$ and (c) for $f(r)$. A portion of the matrix and all diagonal elements are given. (a) Empirical smooth weight function commonly assumed. (b) Weight function derived by the method described in Appendix II. (c) The matrix Π_f given in (A-27) of Appendix III.

the weight matrix \mathbf{P} . In order to avoid such a complicated analysis, it is important to select an optimum interval Δq , by which the analysis can be carried

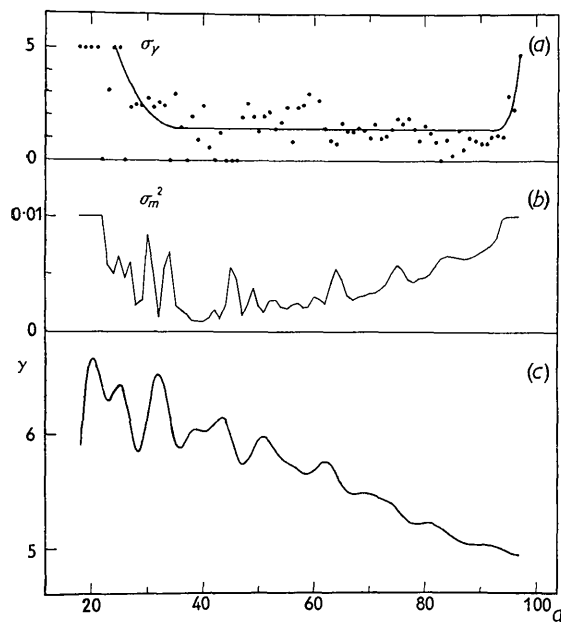


Fig. 2. Estimation of the standard error σ_m^2 for $qM(q)$. (a) Standard error σ_y of the measurement of the y coordinates in the photometer chart. The unit of σ_y corresponds to the minimum scale division at present in use in our laboratory for measuring the y coordinate of the chart. See Appendix II for further significance. Non-essential irregularities which appear in the experimental σ_y (shown in dots) are removed by drawing a smooth curve across the fluctuations, as described in Appendix II. (b) A plausible estimate of σ_m^2 obtained from (a) by the use of equation (A-20) of Appendix II. Conspicuous jags arise from the contributions of σ_x in the regions where the slope of the original y curve (shown in (c)) is steep.

out with a diagonal weight matrix. The consideration of the correlation of observations which leads to a method for estimating a proper interval will be discussed in our forthcoming paper (Murata & Morino, to be published), where a diagonal weight matrix \mathbf{P} corresponding to the observations made at the interval $\Delta q=1$ will be shown to be nearly optimum.

Random errors of measurement on the microphotometer curve originate on one hand from the error Δy in the measurement of the ordinate of the recorder chart, and on the other from the failure Δx in the assignment of the correct abscissa to a pre-determined q scale. Errors which originate from both Δx and Δy are included in the standard error σ_m of $qM(q)$. Even if Δx and Δy are both constant everywhere on the chart, their contributions to σ_m vary from place to place in a manner dependent on the slope of the photometer curve, as stated in Appendix II. It may be shown, therefore, that the diagonal elements of Π_m , which reflect the reliability of our observations on $qM(q)$, do not lie on a uniform function of q , but they fluctuate with a period similar to that of y as illustrated in Figs. 1(b) and 2(c). A reasonable weight matrix Π_m can be estimated from experimental σ_m by eliminating non-essential irregularities. The process is given in Appendix II. Fig. 2 shows a typical example of the dependence of σ_m^2 on q taken from our recent work on silicon tetrachloride (Morino & Murata, 1965).

The weight matrix Π_f for the radial distribution function, $f(r)$, on the other hand, has off-diagonal elements, which indicate that any two 'indirect' observations on $f(r)$, say $f(r)$ and $f(r+\Delta r)$, may be

Table 1. Dependence of least-squares outputs (most probable values and standard errors) on the choice of weight functions (for silicon tetrachloride, in Å)

(a) Least-squares analysis on $qM(q)$ *

	$r(\text{Si-Cl})$	$r(\text{Cl-Cl})$	$l(\text{Si-Cl})$	$l(\text{Cl-Cl})$
A	2.0180 ± 0.0004	3.2932 ± 0.0008	0.0488 ± 0.0007	0.0894 ± 0.0011
B	2.0178 ± 0.0005	3.2936 ± 0.0010	0.0478 ± 0.0009	0.0888 ± 0.0014
C	2.0177 ± 0.0005	3.2950 ± 0.0010	0.0468 ± 0.0010	0.0884 ± 0.0014

* Derived by using the weight functions A, B, and C.

A: Obtained by the process described in Appendix II (Fig. 1(b)).

B: The same as A, but the fluctuations of A are arbitrarily smoothed.

C: An empirical weight function assumed *a priori* (Fig. 1(a)).

(b) Least-squares analysis on $f(r)$ † ($r=0.05 \sim 4.00$ Å; $\Delta r=0.05$ Å)

	$r(\text{Si-Cl})$	$r(\text{Cl-Cl})$	$l(\text{Si-Cl})$	$l(\text{Cl-Cl})$
I	2.0175 ± 0.0005	3.2927 ± 0.0010	0.0474 ± 0.0010	0.0887 ± 0.0013
II	2.0173 ± 0.0007	3.2957 ± 0.0010	0.0459 ± 0.0013	0.0887 ± 0.0013
III	2.0173 ± 0.0010	3.2957 ± 0.0009	0.0459 ± 0.0019	0.0891 ± 0.0011
IV	2.0173 ± 0.0006	3.2957 ± 0.0012	0.0459 ± 0.0012	0.0885 ± 0.0015

† Analyzed by using the weight matrices I, II, III, and IV.

I: A non-diagonal weight matrix given in equation (A-27) and Fig. 1(c).

II: The diagonal unit matrix ($P_{ij} = \delta_{ij}$).

III: A diagonal matrix derived by the law of the propagation of errors: $[(\mathbf{GP}^{-1}\mathbf{G}^*)_{ii}]^{-1}$. ($P_{ij} = \delta_{ij}/r_i$).

IV: The diagonal matrix employed by Bonham & Bartell (1959) ($P_{ij} = r_i^2 \delta_{ij}$).

'dependent'. This situation stems from the fact that both $f(r)$ and $f(r+\Delta r)$ are functions of *all* observations on $qM(q)$. For the example given in Appendix III, the matrix Π_f has a simple form as illustrated in Fig. 1(c). In that particular case, the ratio of the off-diagonal elements for the nearest neighbors to the diagonal is $-\frac{1}{2}$. Although Π_f may in general have a more complicated form, it is readily calculated from equation (5) by using an electronic computer. It is likely that the correlation between $f(r)$ and $f(r+\Delta r)$ fades out rapidly as the interval Δr increases, but the off-diagonal elements among the nearest neighbors cannot be ignored.

We ordinarily apply the least-squares method to the molecular intensity curve rather than to the radial distribution curve, because in the former procedure it is not necessary to take into account the molecular intensity in the region of q where no experimental data are available (*e.g.* $q \lesssim 10$ and $q \gtrsim 100$), and because the weight matrix can be taken as diagonal if observations are made at a proper interval.

It is pertinent to examine how the outputs of our least-squares analysis (most probable values and their standard errors) are dependent on the change in the weight functions. It may be seen from the example given in Table 1 that in this case the outputs are sensitive to the choice of the weights only by as much as their standard errors, and the use of a conventional weight function (Fig. 1(a)) is seemingly justified. The situation, however, depends in general on the quality of the experiment, and the importance of estimating a proper weight function should not be overlooked, particularly if one is concerned with a small difference in the parameters such as a shrinkage effect (Morino, Nakamura & Moore, 1962; Morino, Cyvin, Kuchitsu & Iijima, 1962).

Parameters to be determined

It is well known that to a good approximation the molecular intensity curve has the following expression (Bartell, 1955; Morino, Nakamura & Iijima, 1960), if the influence of the relative phase shift $\Delta\eta$ can be included in the effective parameter l ,

$$qM(q) = \sum_i k_i (A_i/r_i) \exp(-\pi^2 l_i^2 q^2/200) \times \sin(\pi q/10)(r_i - \pi^2 \kappa_i q^2/100). \quad (6)$$

Accordingly, variable parameters to be determined in the analysis are the distances r_i (the r_a parameters (Morino & Iijima, 1962) which are essentially equal to Bartell's $r_g(1)$ (1955)), the mean amplitudes l_i , the anharmonicity parameters κ_i and the indices of resolution k_i . It is seen from equation (6) that the parameters r_i are correlated strongly with κ_i , and also l_i with k_i . Since the parameters κ_i for non-hydrogen distances are generally small, it often happens that they cannot be determined in a statistical sense (Morino, Nakamura & Iijima, 1960; Morino &

Iijima, 1962). In view of the heavy interdependence between r_i and κ_i , however, the latter should not simply be set equal to zero even for non-hydride molecules. Instead, κ_i must be included at least as a finite constant, the order of which may be reasonably estimated by approximate theoretical considerations (Bartell, Kuchitsu & deNeui, 1961; Kuchitsu & Bartell, 1961; Morino & Iijima, 1962).

It seems better to take each index k_i as an independent variable parameter than to assume that they are equal for all atom pairs to be determined. In many cases, however, all indices are shown to be nearly equal, with standard errors for l_i and k_i somewhat larger than those derived when they are all set equal,* as shown in a typical example of our results of silicon tetrachloride listed in Table 2.

Table 2. *Dependence of least-squares outputs on the choice of variable parameters*
(SiCl₄; r and l in Å)*

Parameters†	$r(\text{Si-Cl})$	$r(\text{Cl-Cl})$	$l(\text{Si-Cl})$
A a, c	2.0180 ± 0.0004	3.2932 ± 0.0008	0.0488 ± 0.0007
D b, c	2.0180 ± 0.0004	3.2932 ± 0.0008	0.0485 ± 0.0007
E a, d	2.0176 ± 0.0004	3.2916 ± 0.0008	0.0488 ± 0.0008
F b, d	2.0176 ± 0.0004	3.2916 ± 0.0008	0.0485 ± 0.0007
G a, e	2.0188 ± 0.0008	3.2955 ± 0.0016	0.0487 ± 0.0007

	$l(\text{Cl-Cl})$	$k(\text{Si-Cl})$	$k(\text{Cl-Cl})$
A	0.0894 ± 0.0011	1.081 ± 0.011	1.061 ± 0.018
D	0.0902 ± 0.0007		1.076 ± 0.009
E	0.0894 ± 0.0011	1.082 ± 0.012	1.060 ± 0.019
F	0.0902 ± 0.0008		1.076 ± 0.010
G	0.0895 ± 0.0011	1.081 ± 0.011	1.063 ± 0.018

* Based on the weight function A given in Fig. 1(b).

† a : Indices of resolution k_i varied independently.

b : Both indices taken to be equal and varied.

c : Anharmonicity parameters κ_i taken as constant by a crude theoretical estimation: $\kappa(\text{Si-Cl}) = 1.2 \times 10^{-6}$, and $\kappa(\text{Cl-Cl}) = 7.1 \times 10^{-6}$ in Å³ units.

d : κ_i assumed to be zero.

e : κ_i taken as variable parameters, the outputs being: $\kappa(\text{Si-Cl}) = (3.3 \pm 2.0) \times 10^{-6}$, and $\kappa(\text{Cl-Cl}) = (17.3 \pm 6.3) \times 10^{-6}$.

The discrepancy between $\kappa(\text{Cl-Cl})$ and that estimated in c is not to be taken as very significant. The uncertainty of the distances caused by the error in the estimation of κ is included in the standard error σ (systematic) given in Table 4.

Estimation of random errors

A primary basis for estimating random errors is given by the standard error in equation (4), which we designate as σ_1 . It depends on the discrepancy between our observations and the theoretical molecular intensity curve (6) for which an optimized set of parameters

* Care should be taken because, for one of the two variable parameters which are closely correlated with each other, an improperly smaller standard error is obtained when the other is taken as constant than when it is varied. (Anderson (1958), as cited by Geller (1961)). The situation is shown in Table 2 for r and κ , and for l and k .

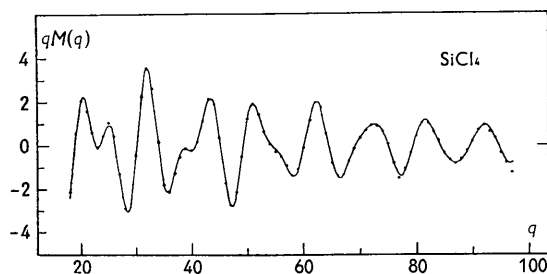


Fig. 3. Least-squares fit on $qM(q)$ for silicon tetrachloride. Solid curve represents the theoretical molecular intensity for which the sum of the weighted square deviations from the observations (shown in dots) is minimized.

are chosen, as illustrated in Fig. 3. While σ_1 is a useful indicator of the quality of the observations, it does not necessarily reveal all non-systematic errors. As another useful measure of the random errors, we may consider a standard deviation, σ_2 , among the most probable values of the parameters obtained from a number of independent sets of observations, *i.e.* from N photographic plates which are taken under as nearly identical experimental conditions as possible and are analyzed individually by the standard process mentioned above. The measures of the random errors, σ_1 (the internal consistency) and σ_2 (the reproducibility), should be of nearly equal magnitude, if the number N is sufficiently large, and if there is no systematic deviation among the plates so that all observations on N plates belong to a single statistical population.* For the mean amplitudes l_i , we usually obtain σ_2 which is of the same magnitude as σ_1 (see Table 4). It sometimes happens, however, that the σ_2 for the distance parameters r_i is somewhat larger than σ_1 , because systematic errors (such as the error of the scale factor, and the error characteristic of any one of the photographic plates), which may not always be eliminated and which are not included in σ_1 , contribute to σ_2 . In such a case, an analysis based on an average curve of N intensities will not lead to a correct estimation of the parameters and their errors; it is therefore important to evaluate σ_2 by a set of individual analyses of N individual intensity curves.

By taking σ_2 , one can find out on the basis of the discrepancy between σ_1 and σ_2 whether some hidden systematic errors are present in the observations, and whether the estimation of σ_1 has any imperfection that makes σ_1 of less statistical significance, such as the use of an improper interval for the diagonal weight function. The estimates σ_1 and σ_2 for the random errors can be regarded as adequate if they turn out to be nearly equal, but if not, it seems appropriate to take whichever is larger as a standard

* For the observations on N sets of plates, the σ_1 for the most probable values of the parameters may be taken as $1/\sqrt{N}$ times the individual standard deviation since in general they are all nearly equal.

error σ , after a careful examination of both σ_1 and σ_2 in regard to the origin of the discrepancy mentioned above.

It is also important to include in the estimates of the errors the uncertainties in the corrections for such various systematic errors (Kuchitsu, 1959; Morino & Iijima, 1962) as are listed in Table 3. When the magnitude of the uncertainties of the correction is not known, it is presumed to be about 50%, or sometimes 100%, of the original correction. An example of our estimates for the random errors is given in Table 4.

Table 3. *Principal origins of systematic errors*

Experimental:	Drift of the accelerating voltage
	Errors in the measurement of the wavelength
	Uncertainty in the camera length
	Imperfections in the sector shape
	Finite sample size
	Irregularity in the q scale of the photometer chart
	Irregularity in the experimental background function
Theoretical:	Failure of the Born approximation
	Uncertainties in the elastic and inelastic scattering factors
	Effect of anharmonicity

Table 4. *Estimation of the errors in the parameters* (SiCl₄; in Å)

	σ_1	σ_2	$\sigma(\text{syst.})^*$	$\sigma(\text{total})$
$r(\text{Si-Cl})$	0.0003	0.0010	0.0009	0.0014
$r(\text{Cl-Cl})$	0.0006	0.0018	0.0017	0.0025
$l(\text{Si-Cl})$	0.0005	0.0006	0.0002	0.0006
$l(\text{Cl-Cl})$	0.0007	0.0009	0.0003	0.0010

* Uncertainties in the corrections for systematic errors. Principal origins of $\sigma(\text{systematic})$ in this case are the following (see Table 3):

Internuclear distances	$r(\text{Si-Cl})$	$r(\text{Cl-Cl})$
Errors in the measurement of the wavelength	0.0004	0.0007
Drift of the accelerating voltage (max. 0.05%)	0.0004	0.0007
Uncertainty in the camera length	0.0006	0.0010
Uncertainty in the parameter κ	0.0004	0.0010
Total	0.0009	0.0017
Mean amplitudes	$l(\text{Si-Cl})$	$l(\text{Cl-Cl})$
Effect of the finite sample size	0.0002	0.0003

In some cases, it is convenient to represent the 'limit of error' as $\pm 2.5\sigma$, which specifies the 99% confidence interval, in the sense that the probability of finding the result outside this range is only less than 1% (Morino & Iijima, 1962).

Concluding remarks

As is well known, an ultimate appraisal of the accuracy of our experimental and statistical method should be made possible by a comparative study of molecules whose structures are exactly known by spectroscopic

studies. Since the theoretical background which correlates the parameters obtained by electron diffraction with the spectroscopic parameters has now been established both for diatomic (Bartell, 1955) and for polyatomic molecules (Morino, Kuchitsu & Oka, 1962), there should be no ambiguity in principle in this comparison. A systematic study for this purpose is now being undertaken in our laboratory, and the result will be published in the near future. For the carbon disulfide molecule which we have studied so far, the agreement of the distances, the mean amplitudes, and the shrinkage effect with the corresponding spectroscopic values within the claimed uncertainties is encouraging, as shown in Table 4 of the paper by Morino & Iijima (1962).

APPENDIX I

Weight matrix for indirect observations

Suppose a set of indirect observations μ is derived from a set of direct observations \mathbf{M} by linear combinations:

$$\mu = \mathbf{G}\mathbf{M}. \quad (A-1)$$

By operating \mathbf{G} on equation (1) of the text and by denoting $\mathbf{G}\mathbf{F}$, $\mathbf{G}\mathbf{V}$, and $\mathbf{G}\mathbf{A}$ by ϕ , ϵ , and α respectively, we get

$$\phi = \mu - \epsilon = \alpha\mathbf{X}, \quad (A-2)$$

which is in a style similar to equation (1). Since the condition of least squares should exist among the residuals for the *direct* observations, the condition should be the same as that given by equation (2),

$$\mathbf{V}^*\mathbf{P}\mathbf{V} = \text{minimum}.$$

Our problem is to find a matrix Π in the expression

$$\epsilon^*\Pi\epsilon = \text{minimum},$$

that is,

$$\alpha^*\Pi\alpha\mathbf{X} = \alpha^*\Pi\mu, \quad (A-3)$$

so that the normal equations in terms of α and μ may be obtained. While \mathbf{G} is in general a non-square matrix, the problem is conveniently handled by means of Lagrange's method of undetermined multipliers. The condition is given by

$$\Psi = \mathbf{V}^*\mathbf{P}\mathbf{V}/2 + \Lambda^*\Phi = \text{minimum}, \quad (A-4)$$

where

$$\Phi = \mu - \epsilon - \alpha\mathbf{X} = 0. \quad (A-5)$$

By taking partial derivatives, it is shown that

$$\partial\Psi/\partial\mathbf{V} = \mathbf{V}^*\mathbf{P} - \Lambda^*\mathbf{G} = 0, \quad (A-6)$$

and accordingly,

$$\Lambda^* = \mathbf{V}^*\mathbf{G}^*(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1} = \epsilon^*(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}, \quad (A-7)$$

and that

$$\partial\Psi/\partial\mathbf{X} = -\Lambda^*\alpha = 0, \quad \text{or} \quad \alpha^*\Lambda = 0. \quad (A-8)$$

Since \mathbf{P} is in general a symmetrical matrix, equations (A-7) and (A-8) lead to

$$\alpha^*(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}\epsilon = 0, \quad (A-9)$$

and hence, the following set of normal equations may be obtained from equation (A-2),

$$\alpha^*(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}\alpha\mathbf{X} = \alpha^*(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}\mu. \quad (A-10)$$

Consequently, the least-squares process for indirect observations can be made analogous to that for direct observations, provided that the weight matrix is taken as

$$\Pi = (\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)^{-1}. \quad (A-11)$$

As may be seen from the derivation of equation (A-7) from (A-6), the number of the observations μ should not exceed that of \mathbf{M} in order that the matrix $(\mathbf{G}\mathbf{P}^{-1}\mathbf{G}^*)$ may be non-singular.

Although our data are represented by non-linear equations, it is well known that the first-order Taylor expansion makes the present argument hold without modification, if the parameters \mathbf{X} , the observations \mathbf{M} or μ , and the matrix \mathbf{G} are regarded as the corrections to the trial parameters, the difference between the observed and the trial values, and the relevant derivative matrix, respectively. It can also be shown that the standard error is given by an equation analogous to equation (4) with corresponding changes in the notation.

The general formula (A-11) covers the following cases† for deriving proper weight matrices Π .

(a) When the column matrices of observations, \mathbf{M} and μ , are related to each other by one-to-one correspondence, *i.e.* when any one of the μ elements, μ_j , depends on only one of the \mathbf{M} elements, M_i , and *vice versa*, the matrix \mathbf{G} may be taken as diagonal; it obviously follows that Π is diagonal if \mathbf{P} is diagonal,

$$\Pi_{jj} = P_{ii}/(G_{ji})^2. \quad (A-12)$$

(b) When any one of the \mathbf{M} elements, M_i , is related to only one of the μ elements, μ_j , whereas one of the μ elements, μ_j , is related to more than one \mathbf{M} element, say $M_i, M_{i+1}, \dots, M_{i+n}$, \mathbf{G} is a rectangular matrix with zero cross elements ($G_{kl} = 0$ for $k = j$ and $l = i, i+1, \dots, i+n$, and for $k = j$ and $l < i$ or $l > i+n$). If \mathbf{P} is diagonal, equation (A-11) leads to a diagonal Π . The procedure is equivalent to that of the least squares where restrictive conditions are present. An example of this case is given in Appendix II, where a method similar to Deming's curve fitting is applied to the measurements of x, y coordinates, in both of which experimental errors are involved (Deming, 1946).

(c) When one of the \mathbf{M} elements is related to more than one μ element, off diagonal elements appear in Π even if \mathbf{P} is diagonal; the observations μ are no longer

† When the matrix Π is diagonal, as mentioned in cases (a) and (b), equation (A-11) represents the ordinary law of propagation of errors.

mutually independent. An example of this case is given in Appendix III, where a non-diagonal weight matrix Π_f for the radial distribution curve $f(r)$ is derived.

APPENDIX II

Estimation of the weight function Π_m for $qM(q)$

In the sector-microphotometer method of gas electron diffraction, one makes direct observations on the x and y values of the photometer record at a certain prescribed interval. With allowance for the errors in the measurements of both x and y coordinates, Δx_i and Δy_i respectively, the least-squares criterion for a curve fitting is given by

$$\Sigma(p_{xi}\Delta x_i^2 + p_{yi}\Delta y_i^2) = \text{minimum.} \quad (A-13)$$

The weight functions p_x and p_y must be so taken that they are inversely proportional to the squared standard errors of the measurement of x and y values respectively. The weight functions p_x and p_y may be estimated with less difficulty than that for $qM(q)$, Π_m , since the former are related to the direct observables x and y , and since they are likely to be smooth and uniform; it is also easier to interpret the physical significance of the p_x and p_y which are determined by an empirical procedure.

Our primary problem is to find the matrix \mathbf{G} which relates $qM(q)$ with x and y . The indirect observations μ on $qM(q)$ are written as the following difference,

$$\mu = qM(q)^{\text{obs}} - qM(q)^{\text{calc}}, \quad (A-14)$$

since this is a non-linear case. The functional dependence of μ on x and y is given by the following relations.

$$qM(q)^{\text{obs}} = q(I_t/I_b - 1) = q\{[F \log_{10}(y_0/y) - I_e]/I_b - 1\}, \quad (A-15)$$

and

$$q(x) = (40/\lambda) \sin \left[\frac{1}{2} \tan^{-1}(x/a_x L) \right], \quad (A-16)$$

where the symbols have the following meaning:

x : the x distance on the photometer trace from the center,

y and y_0 : the y distances of the record and the bright line, respectively, from the dark line on the photometer trace,

I_t , I_b , and I_e : the intensities of the total scattering, the background, and the extraneous scattering respectively,

F : the density-intensity calibration function, taken as unity in the present study (Morino & Iijima, 1962),

$qM(q)^{\text{calc}}$: the molecular intensity calculated by using equation (6) with assumed values of the parameters,

a_x : the ratio of the x scale of the chart to that of the original photographic plate (10 in this case),

L : the camera distance, and

λ : the electron wavelength.

For the μ given in equation (A-14), the matrix \mathbf{G} takes the following expression,*

$$\mathbf{G} = \begin{pmatrix} G_{x_1} & G_{y_1} & 0 & 0 & 0 & 0 & \dots & \dots \\ 0 & 0 & G_{x_2} & G_{y_2} & 0 & 0 & \dots & \dots \\ 0 & 0 & 0 & 0 & G_{x_3} & G_{y_3} & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}, \quad (A-17)$$

where

$$G_x = \frac{dq}{dx} \left[\left(\frac{I_t}{I_b} - 1 \right) - q \frac{I_t}{I_b^2} \frac{dI_b}{dq} - q \frac{1}{I_b} \frac{dI_e}{dq} - \frac{d(qM(q)^{\text{calc}})}{dq} \right], \quad (A-18)$$

$$G_y = (q/I_b a_y) (\partial I_t / \partial y)_x = -q \log_{10} e / I_b a_y y, \quad (A-19)$$

and a_y is the degree of linear expansion of the record in the y direction (about 5 in this case).

According to equations (A-11) through (A-19), the weight matrix Π_m of a diagonal form is obtained,

$$(\Pi_m)_{ii}^{-1} = G_{xi}^2 p_{xi}^{-1} + G_{yi}^2 p_{yi}^{-1}, \quad (A-20)$$

and hence the standard error σ_m of $qM(q)$ is given by

$$\sigma_m^2 = G_x^2 \sigma_x^2 + G_y^2 \sigma_y^2, \quad (A-21)$$

where σ_x and σ_y are the standard errors of the measurement of the x and y values respectively. (See footnote, p. 554.)

The magnitude of G_x is especially sensitive to the slope of the photometer curve, and the contribution from σ_x is important in the region where the y curve is steep. Then σ_m is expected to be larger, and hence, a less weight is to be assigned to the observations of $qM(q)$ in that region. This is the principal origin of the fluctuation of Π_m with q mentioned in the text. An example is given in Fig. 2.

A reasonable weight function Π_m may be estimated by eliminating accidental irregularities from the experimental σ_m by the following steps:

(a) Molecular intensity curves obtained from a number of photographic plates are processed separately by the least-squares method by taking an initial weight function.

(b) The standard error σ_m is assigned to $qM(q)$ for each q by taking the square of the difference between the observed and calculated $qM(q)$ for several plates (*i.e.* the expectation of the square of residuals).

* The background I_b depends on y , since I_b is drawn empirically through the observed y values. For the present problem, however, I_b ought to have no 'explicit' dependence on y . That is, the errors Δy do not interfere with the drawing of the I_b , and therefore $\partial I_b / \partial y = 0$, since the errors Δy fluctuate randomly around the true values, and since I_b should be a smooth function. It is readily shown from equations (A-18) and (A-19) that $G_x = (dy/dx)G_y$; this agrees with the following simple reasoning of the origin of G_x . Suppose the measurement of x is in error by Δx . The measurement of y must then be in error by $(\Delta y)_x = (dy/dx)\Delta x$, which will cause $qM(q)^{\text{obs}}$ to be in error by $G_x \Delta x = G_y (\Delta y)_x = G_y (dy/dx) \Delta x$.

(c) The standard error σ_x , which may be taken as constant through the photometer chart, is estimated by experience.*

(d) By subtracting the contribution of σ_x^2 from σ_m^2 , σ_y^2 is obtained for each q by using equation (A-21).† Since σ_y is considered to be a slowly changing function of q , the σ_y thus derived is made smooth by ignoring non-essential irregularities. (See Fig. 2(a)).

(e) The standard error σ_m is obtained from the smoothed $\sigma_y(q)$ by using the same equation (A-21), as shown in Fig. 2(b), and $\Pi_m(q)$ is taken to be σ_m^{-2} (Fig. 1(b)).

(f) The weight function is made self-consistent, if necessary, by a recycle in which the weight function Π_m obtained in (e) is used in the step (a). Most portions of the above steps can be automatized by the use of an electronic computer.

APPENDIX III

A weight matrix Π_f for the radial distribution curve

Let us take an example in which the weight function Π_m for $qM(q)$ is represented as

$$\Pi_m(q, q') = q^2 \exp(-2bq^2) \delta_{qq'}. \quad (A-22)$$

Since $f(r)$ is given by

$$\begin{aligned} f(r) &= \int_0^\infty qM(q) \exp(-bq^2) \sin(\pi qr/10) dq \\ &\equiv \sum_q qM(q) \exp(-bq^2) \sin(\pi qr/10), \quad (A-23) \end{aligned}$$

the elements of the matrix \mathbf{G} are shown to be

$$\mathbf{G}(r, q) = \exp(-bq^2) \sin(\pi qr/10). \quad (A-24)$$

It follows from equation (5) of the text that

* The standard error σ_x is taken as the minimum scale division s_x (0.1 mm) at present in use in our laboratory for measuring the x coordinate of the photometer chart. (The record is usually subject to a mechanical linear magnification of ten times the x scale on the original photographic plate.) By changing the σ_x values, it is found that the feature of jags which appear in the resulting σ_m^2 depends on σ_x ; the output parameters of the least-squares analysis, however, are not sensitive to the choice of σ_x . The σ_x taken as s_x is regarded as reasonable in this case since the peak values of the resulting σ_m^2 in the regions where the effect of σ_x is large are nearly equal to those of the observed σ_m^2 curve. One could expect an ideal σ_x to be about $s_x/2$; this, however, results in a somewhat worse correspondence with the peak values of the observed σ_m^2 indicating that the presence of other random errors makes the σ_x to be nearly equal to s_x .

† For the example given in Fig. 2, the ordinate scale unit of σ_y is taken as the minimum scale division (1/1000 of the chart full scale, 24 cm) at present in use in our laboratory for measuring the y coordinate of the chart. (The record is usually subject to an electronic linear magnification of about five times the original y scale.) The random experimental error of the measurement σ_y is therefore shown to be comparable with the uncertainty associated with the minimum scale division.

$$\begin{aligned} \Pi_f^{-1}(r, r') &= [\mathbf{G}\Pi_m^{-1}\mathbf{G}^*](r, r') \\ &\equiv \int_0^\infty q^{-2} \sin(\pi qr/10) \sin(\pi qr'/10) dq \\ &= (\pi^2/20) \min(r, r'). \quad (A-25) \end{aligned}$$

The matrix Π_f^{-1} is therefore

$$\Pi_f^{-1} = \pi^2/20 \begin{pmatrix} r_1 & r_1 & r_1 & r_1 & r_1 & \dots \\ r_1 & r_2 & r_2 & r_2 & r_2 & \dots \\ r_1 & r_2 & r_3 & r_3 & r_3 & \dots \\ r_1 & r_2 & r_3 & r_4 & r_4 & \dots \\ r_1 & r_2 & r_3 & r_4 & r_5 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix}. \quad (A-26)$$

By denoting the difference $r_{n+1} - r_n$ as Δr and taking the inverse, it is shown that the weight matrix Π_f is given by the following simple form

$$\Pi_f = \frac{20}{\pi^2 \Delta r} \begin{pmatrix} r_2/r_1 & -1 & 0 & 0 & \dots & \dots & \dots \\ -1 & 2 & -1 & 0 & \dots & \dots & \dots \\ 0 & -1 & 2 & -1 & \dots & \dots & \dots \\ 0 & 0 & -1 & 2 & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & 0 & -1 & 2 & -1 \\ \dots & \dots & \dots & \dots & 0 & 0 & -1 & 1 \end{pmatrix}. \quad (A-27)$$

In this case, all diagonal elements are equal except for those at both ends, the differences being trivial for the present consideration. Thus the ratio of the off-diagonal elements for the nearest neighbors to the diagonal is $-\frac{1}{2}$, and all other off-diagonals are zero.

The authors are indebted to Prof. J. Kakinoki and Drs T. Ino and K. Katada of Osaka City University for their valuable criticism and advice. They are also indebted to the Ministry of Education of Japan for the research grant.

References

- ANDERSON, T. W. (1958). *An Introduction to Multivariate Statistical Analysis*. New York: Wiley.
 BARTELL, L. S. (1955). *J. Chem. Phys.* **23**, 1219.
 BARTELL, L. S., KUCHITSU, K. & DE NEUI, R. J. (1961). *J. Chem. Phys.* **35**, 1211.
 BASTIANSEN, O., HEDBERG, L. & HEDBERG, K. (1957). *J. Chem. Phys.* **27**, 1311.
 BONHAM, R. A. & BARTELL, L. S. (1959). *J. Chem. Phys.* **31**, 702.
 DEMING, W. E. (1946). *Statistical Adjustment of Data*. New York: Wiley.
 GELLER, S. (1961). *Acta Cryst.* **14**, 1026.
 HAMILTON, W. C. (1954). Thesis, California Institute of Technology.
 HEDBERG, K. & IWASAKI, M. (1962). *J. Chem. Phys.* **36**, 589.
 HEDBERG, K. & IWASAKI, M. (1964). *Acta Cryst.* **17**, 529.
 KUCHITSU, K. (1959). *Bull. Chem. Soc. Japan*, **32**, 748.

- KUCHITSU, K. & BARTELL, L. S. (1961). *J. Chem. Phys.* **35**, 1945.
- MORINO, Y., CYVIN, S. J., KUCHITSU, K. & IJIMA, T. (1962). *J. Chem. Phys.* **36**, 1109.
- MORINO, Y. & IJIMA, T. (1962). *Bull. Chem. Soc. Japan*, **35**, 1661.
- MORINO, Y., KUCHITSU, K., IJIMA, T. & MURATA, Y. (1962). *J. Chem. Soc. Japan (Pure Chem. Sec.)*, **83**, 803.
- MORINO, Y., KUCHITSU, K. & OKA, T. (1962). *J. Chem. Phys.* **36**, 1108.
- MORINO, Y. & MURATA, Y. (1965). *Bull. Chem. Soc. Japan*, **38**, 104.
- MORINO, Y., NAKAMURA, Y. & IJIMA, T. (1960). *J. Chem. Phys.* **32**, 645.
- MORINO, Y., NAKAMURA, J. & MOORE, P. W. (1962). *J. Chem. Phys.* **36**, 1050.
- Proc. Intern. Conf. Mag. Cryst. II. Electron and Neutron Diffraction* (1962). *J. Phys. Soc. Japan*, **17**, Suppl. B-II.

Short Communications

Contributions intended for publication under this heading should be expressly so marked; they should not exceed about 1000 words; they should be forwarded in the usual way to the appropriate Co-editor; they will be published as speedily as possible. Publication will be quicker if the contributions are without illustrations.

Acta Cryst. (1965). **18**, 557

The space group of LaAlO₃ below 720 °K, and the compounds NdAlO₃ and PrAlO₃, probably isomorphous at room temperature. By B. DERIGHETTI, J. E. DRUMHELLER, F. LAVES, K. A. MÜLLER and F. WALDNER, *Physik-Institut der Universität Zürich, Institut für Kristallographie und Petrographie der ETH Zürich, and IBM Zürich, Research Laboratory Rüschlikon, Switzerland*

(Received 20 August 1964)

Geller & Bala (1956) investigated LaAlO₃ by X-rays and found a symmetry change at 720 °K (independently confirming results of Gränicher & Müller, 1957). Whereas LaAlO₃ has the ideal perovskite structure (*i.e.* cubic with one formula unit per cell) above this temperature, a rhombohedral deformation takes place at 720 °K continuously increasing with decreasing temperature. In addition, extra reflexions observed at room temperature indicated a doubling of the rhombohedral *c* axis. Geller & Bala proposed as most probable the space group $R\bar{3}m(D_{3d}^5)$ with La in $(c) - 3m \pm (0, 0, z)$, Al in $(a) - \bar{3}m (0, 0, 0)$ and $(b) - \bar{3}m (0, 0, \frac{1}{2})$, and O in $(h) - m \pm (x\bar{x}z; x, 2x, z; 2\bar{x}, \bar{x}, z)$ if hexagonal axes are chosen. On the other hand, consistent with their published X-ray data the space group $R\bar{3}c(D_{3d}^5)$ with La in $(a) - 32 \pm (0, 0, \frac{1}{2})$, Al in $(b) - \bar{3} (0, 0, 0; 0, 0, \frac{1}{2})$ and O in $(e) - 2 \pm (x, 0, \frac{1}{4}; 0, x, \frac{1}{4}; \bar{x}, \bar{x}, \frac{1}{4})$ would also be possible. In $R\bar{3}m$ three parameters would be needed to characterize the structure; however, only one would be needed in $R\bar{3}c$, and the Al positions would be equivalent.

The nuclear quadrupole resonance measurements of ¹³⁹La and ²⁷Al (Müller, Brun, Derighetti, Drumheller & Waldner, 1964) and the electron paramagnetic measurements of crystals doped with Gd³⁺, Cr³⁺ (Kiro, Low & Zusman, 1963) and Fe³⁺ (Müller *et al.*, 1964) give information on the point symmetry of the La and Al positions and indicate that the space group $R\bar{3}m$ is improbable. From the data on ¹³⁹La and Gd³⁺ one sees one La site, its surroundings being axially deformed. The data on ²⁷Al and Cr³⁺ show also only one Al site with axially deformed surroundings. The electron paramagnetic resonance spectrum of Fe³⁺ consists of two sets of lines showing two sets of cubic axes rotated along the rhombohedral *c* axis, whereas the distances to the nearest neighbours seem to be equal for both sets. The rotation angle $\pm\alpha$ is measured to be $6.0^\circ \pm 0.2^\circ$ at $298^\circ \pm 3^\circ$ K.

Thus, a structure with equivalent La sites and equivalent Al sites would be more consistent with the measurements mentioned above. Only two space groups, $R\bar{3}c$ and $R\bar{3}c$, would conform with the X-ray data and the resonance data. Further, the observed two sets in the Fe³⁺ spectrum could be attributed to the two types of equivalent Al positions which would have in both space groups rotated near-octahedral surroundings related by glide planes. As there are no data yet known to necessitate the assumption of $R\bar{3}c$ we propose $R\bar{3}c$ as the most probable space group for LaAlO₃.

The oxygen parameter $x \neq \frac{1}{2}$ is still unknown but should be strongly temperature dependent. A calculation from the angle α of the Fe³⁺ spectrum would give $x = 0.53$ at room temperature in hexagonal coordinates, if one neglects possible structure deformations by the impurity ion.

From a crystal-chemistry point of view it is interesting to note that the more x deviates from $\frac{1}{2}$ the more the La coordination number deviates from 12 (in the ideal perovskite structure) to 9. As the compounds NdAlO₃ and PrAlO₃ are reported to be isomorphous with LaAlO₃ (Geller & Bala, 1956) it is probable that they too have the space group $R\bar{3}c$.

The authors are indebted to R. Jansen and F. Forrat for preparing the Fe³⁺-doped LaAlO₃ crystal.

References

- GELLER, S. & BALA, V. B. (1956). *Acta Cryst.* **9**, 1019.
- GRÄNICHER, H. & MÜLLER, K. A. (1957). *Nuovo Cim. Suppl.* **6**, Ser. X, 1216.
- KIRO, D., LOW, W. & ZUSMAN, A. (1963). *Proc. First Intern. Conf. Paramagnetic Resonance. Vol. I.*, 44. New York: Academic Press.
- MÜLLER, K. A., BRUN, E., DERIGHETTI, B., DRUMHELLER, J. E. & WALDNER, F. (1964). *Phys. Letters* **9**, 223.